

Broadly stable atmospheric CO₂ and CH₄ levels over the past 3 million years

Corresponding Author: Ms Julia Marks-Peterson

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Attachments originally included by the reviewers as part of their assessment can be found at the end of this file.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)

This manuscript builds on data obtained from new ice cores drilled in the Allan Hills region enabling to get snapshots of old atmosphere and hence to infer the evolution of CO₂ and CH₄ in the atmosphere over the last 3 millions years. The main finding is that the CH₄ level did not significantly change over the last 2.9 millions years and that CO₂ level only slightly decreased by approximatively 25 ppm.

This study complement the study of Yan et al. (2019) by adding more core and significantly more dated ice level. The conclusion is also a bit different. While Yan et al. (2019) commented on the reduced amplitude of the amplitudes of the glacial – interglacial cycles for both climate and greenhouse gases concentration, the present manuscript only comment on the « mean » level (which may actually be biased toward interglacial values).

This study has a broad enough impication to be of high interest for publication in Nature. Still, there are many points requiring further investigations or discussions in order to prevent misinterpretation.

1- The authors did some efforts in the companion paper to explain the different drilling also including a map in the supplementary information. Still, it could be nice to have it here also with a clear indication of what has been measured in each cores used here for the synthesis (ALHIC 1502, 1503, 1901, 1902, 1903), especially the age – depth relationship for each core. This could also help to understand why the conclusion from this manuscript is different from the conclusion of the paper of Yan et al., 2019. The differences in interpretation (especially about the amplitude of the glacial – interglacial cycles) should also be explained in the mani manuscript to help following the research on these ice cores network.

2- I am missing a reconstruction of the age for the different ice level from the measurements of the d40/38 and d38/36. The data are never shown nor the limitations associated with the method (and discussed in other previous publications such as Yau et al.,). However, when looking at the data provided as a table, we see that ages of several millions years associated with strange d38/36 and d40/36 data were removed but there is no explanation on why / how there was a selection in the data shown here. Recent communications have shown that 81Kr dating can be performed on samples of 1-1.5 kg. It would b every helpful knowing the limitations with the 40Ar,atm dating to provide a few 81Kr dates.

3- When looking at the d38/36 data in the supplementary, there is no link between these data and the d18Oice. This is a bit surprising since there should be a link because of gravitationnal fractionation. Have you investigated it ? Any explanation ? Also the d38/36 data for the top part are quite high compared to the rest. Again any explanation for these data ? May it influence the 40Ar,atm dating ? The d38/36 data should be discussed a bit in the supplement at least (either here or in the comanion paper).

4- I am surprised by the assumption that CO₂ and CH₄ « values agree well with depth-weighted averages of these gases from the same period in EDC (sup figure S3) ». Indeed, on his figure, there is as important proportin of the red stars falling on

high level of CO₂ and mid-level of CH₄ (such CH₄-CO₂ characteristics are rare in the EDC records). This is surprising and should be explained because it is also linked to the main conclusion of the manuscript stating that CO₂ levels are a bit higher than pre-industrial levels for old period while CH₄ levels are similar to pre-industrial level.

5- The study related to figure S2 should be improved for two reasons. First, it seems that the authors assumed the same diffusion for CO₂, CH₄, d18O_{atm} while the diffusivities and solubilities are probably different (see Bereiter et al., 2009). This should be calculated / discussed. Second, I am not really sure how you want to extract interglacial values for d18O_{atm}. The beginning of the interglacial periods (according to CO₂ – CH₄) show high d18O_{atm} (1 permil or so) and then d18O_{atm} decreases. The initial distribution of d18O_{atm} is also different from the initial distributions of CO₂ and CH₄ with very few high d18O_{atm} values – this could also affect the comparison that you are discussing.

6- Figure 1 – I am unable to see the « open blue circles » referred to in the caption

7- I am surprised by the mention of negative Ar/N₂ ratio ? This is not possible ? You mean significantly low dAr/N₂ ? Could you provide some more explanations in the text by giving values. Again, there is a lack in showing the elemental and isotopic values for argon.

8- The correction for CO₂ levels using d13C_{atm} is simple but I agree that the authors can not really do better. It is important to show on the main figure 1 that you have different types of data. Also this biological effect can probably also affect CH₄ and d18O_{atm} and was also taken into account ? Is it possible to discuss this a bit or are the samples removed from the analysis ?

9- There are some strange values in the ALHIC1503 series display on supplementary figure 1 : very low dO₂/N₂ data, very low d15N – what was done with these data ? I imagine that they are not further used here because the ice is not dated but it would be nice to comment since you are showing them.

Referee #2

(Remarks to the Author)

This paper presents a new dataset of CO₂ and CH₄ values from blue ice cores from Allan Hills, Antarctica. The novelty is that the data extend back to 3 Ma, well beyond the conventional ice core time period. Compared to an earlier paper (Yan et al 2019) that extended 2 Ma in the past, there is a greatly improved dating, allowing an attempt to define trends with time, and the ability to compare with datasets of CO₂ from marine proxies and from models.

If it can be taken at face value, the results shown here (notably no change in CO₂ over the MPT, and a slow decline between 3.0 and 1.2 Ma) are very exciting and important. Given that they come from the more direct measurement of air in ice, they would effectively supersede and arbitrate between earlier compilations from different marine methods. However there are some issues that are treated in the supplement but glossed over in the text that might cause us to question what we are actually seeing here and whether it is an objectively superior datasets to the others. I don't think these preclude publication of the data, but they do require the authors to be rather clearer about the different options the dataset might represent and they make it less of a clear breakthrough than might be implied on a straight reading of Figure 1.

The issue is what kind of an average each datapoint represents and whether that average has a meaning that is constant as we go back in time. This is covered cursorily in lines 113-136 of the main paper, and more extensively in lines 90-164 of the supplement. The issue arises from the fact that only long-term averages can be derived in the older samples, and we would ideally like them to be reliable time averages of periods around the 40Ar age estimate. However such mixed samples can suffer biases from two causes. Firstly there is more snowfall in warm than cold periods, so we expect averages that cover glacial cycles to be biased towards interglacial values. Secondly there are clearly flow disturbances in the ice, and experience in Greenland and Antarctica show that this can cause distortion of the age scale particularly around the interface between glacial and interglacial ice (due to their different rheology). This second issue might in theory lead to biases either towards glacial or interglacial values.

The only direct evidence the authors have about this issue is in the ice they refer to as “short snapshots”, which I believe is the ice above 141 m in ALHIC1901 (red in Fig S3B). There they observe a clear interglacial bias (using EDC as reference), which they ascribe to an extreme accumulation rate contrast between interglacials and glacials at Allan Hills. This explanation (also partly justified in Fig S2) seems reasonable. But then when they reach the deeper ice (146-148 m and ~151 m) of similar age, they find values (in a range of parameters) closer to the average at EDC. This is curious, and the authors use it to justify an assertion that all the values below 141 m are similar to time averages. This is really quite a strong assumption on very weak evidence and poses a number of challenges:

- a) If there is a strong interglacial bias in accumulation rate it must surely apply to the area of origin of all the Allan Hills ice, so why is the bias seen above 141 m not present below 141 m? The authors talk about a weaker accumulation contrast in older ice (line 157-160 of supplement), but this doesn't explain the apparent reduced bias in ice younger than 800 ka below 141 m. (And even a 3:1 contrast as opposed to a 10:1 contrast in accumulation would impose a significant bias.)
- b) The most obvious answer to (a) is that the deeper ice also suffers a flow disturbance that has prioritised glacial ice, and therefore partly restored the balance between glacial and interglacial material.
- c) If a and b are both in play then what is the justification for assuming that the older ice (142-146 m, 148-151 m, and below 152 m) has the same bias as the ice between 146 and 148 m?

d) A final issue, but perhaps less fundamental, is that the ice before the MPT (with what in benthic isotopes looks like symmetric glacial/interglacial cycles) imposes a fundamentally different averaging compared to the sawtooth shaped cycles of the post-MPT, which if there was no accumulation bias would have an average tending towards glacial values.

The result of these issues is that it raises the fundamental question whether the averages have the same meaning as we progress back in time, which then raises questions about the trends. At one extreme, one could assume that the accumulation rate bias seen in the short snapshots applies to older ice, and that the long snapshots of younger ice below 141 m are anomalous. This would lead to a very uncomfortable conclusion about interglacial concentrations but cannot be entirely discounted. At the other extreme one can assume that somehow (though the paper doesn't explain how) the accumulation rate bias is completely lost in the long snapshots younger than 800 ka, and that this same correction of the bias applies throughout the ice below 141 m (and its equivalent in ALHIC1503). This also seems implausible to me, but is the interpretation the authors have adopted. It leads to the also rather uncomfortable conclusion that the higher boron values, including some around 1.2 Ma (Fig S6, pink and purple averages), are wrong.

More realistically one could propose that in the older ice the values are closer to an average (due to a smaller accumulation contrast, and likely a weaker rheology contrast), but could contain a bias towards either glacial or interglacial values, imposing an uncertainty of probably tens of ppm on the numbers. Since this bias might vary systematically with depth and age, this will make it hard to define the trend from 3 Ma, or confirm the lack of trend across the MPT. Possibly that weakening of the message makes this a paper that warrants publication somewhere other than Nature, but that would be an editorial decision.

My recommendation therefore is that the authors should discuss in the main paper these alternative options and attempt to add a realistic preservation uncertainty to their data in order to provide a time series with appropriate uncertainty.

Because my issues with the paper are quite fundamental and need addressing before it proceeds further I add only a few less important technical issues below. I should add that the actual analytical achievement is impressive and I have no complaints or concerns about the analytical values or ages used in this paper.

Minor technical issues

Line 83. Shackleton submitted. I am assuming this is not the parallel MOT paper but rather a related paper about Miocene and Pliocene ice. I understood that Nature doesn't allow unpublished work to be referenced so this will have to depend whether that other paper is published before this one.

Short and long snapshots. It would have been very helpful if the data table (xls file) had a column saying which samples are used as long and which as short snapshots. This could also have been more clearly defined in the text.

Figure 1. I believe this is only the data that have NOT been corrected using $\delta^{13}C$. Please clarify this in the caption. I have to say that the correction, while clever and interesting, is a bit of a distraction, since none of the conclusions of this paper rely on any of the corrected data. I would recommend that the authors use the main text to explain their ideas about averaging properly and take most of the discussion of correcting CO_2 -rich data into the supplement or a separate paper.

Line 156 "our CO_2 concentrations across the MPT agree with the CO_2 -leafwax record". This is way too strong a statement based on Fig S7A. Sure the concentrations are similar but the the long term smoothing of the leafwax data contains a pattern I can't see in the ice core data.

Supplement line 91-93, do these percentage variances refer to only the short snapshots? Please clarify.

Version 2:

Reviewer comments:

Referee #1

(Remarks to the Author)

Many thanks for this second version of the manuscript. I now understand that the dating discussion is in another manuscript which is available, it helps a lot.

These unique data are very important and should be published. I still have some questions on the method and I found that there was a lack of information some time. I detail the comments below.

1- Summary paragraph :

Why do you speak about « secular changes » here ? Why do you use the adjective « secular » ?

2- L. 81 « individual folds and age discontinuities are resolved » -> Are we 100% sure of it ? How can we exclude that some folds or hiatus can still exist at smaller scale ?

3- L. 103 – I do not really understand the concept of « age density » . We can not exclude hiatus in the record, especially below 140 m so how would age density be defined in this case ? Also it is never mentioned that sections of ice are missing in this ice core. It is always about thinning and folding. Some thoughts about hiatus should be introduced.

4- Related to the previous comment, it is important to understand what you mean by « climate dependent preservation bias ». Is it that you have more hiatus during certain period ? Or more diffusion / thinning during a certain climate ?

5- The term « physical mixing » is not clear. What do you mean ? A mixing of two layers of ice because of very small scale folding ? Or diffusion of the tracers within the ice among adjacent layers ? Please further explain what physical process you have in mind.

6- From l. 126 (and despite the supplement), it lacks explanations and more description with additional figures (can be in the supplement).

a. I can not understand how you create a synthetic 400-800 ka ice core record using glacial – interglacial thickness variations from the text and figure S2 which is not very useful as it is (we do not see the glacial – interglacial cycle). I suggest to simply show the synthetic curve and explicitly give the formula used to produce it.

b. « Thinning and small-scale folding were simulated by applying a gaussian smoothing. » So if I understand well, you do not apply ice diffusion ? Still, ice diffusion is very important for ice older than 1 million year and should be considered (Bereiter et al., 2014). This effect should be implemented.

c. « Applying a gaussian smoothing » -> with what parameters ? Please give the detail of this gaussian smoothing (standard deviation of the distribution). It is important that the details of the calculations are given to follow what is done.

d. I am surprised by the sentence beginning l. 145. The diffusion for O₂, CO₂ are given in the paper by Bereiter et al. (2014) so it is possible to infer the effect of diffusion for both CO₂ and d18O_{atm}.

7- I am not fully convinced by the d38/36Ar argument against gas loss (l. 156). As d18O_{ice} is varying with time, d38/36Ar could also vary with time because of changes in accumulation rate and temperature. In parallel, you mention O₂/N₂ measurements in the supplement which you could use for this purpose (and you actually used for correction of the gas loss effect – so there is some gas loss effect ? This is not clear). Why not showing and using them ?

8- Figure 2 –

a. there is a significant difference between the yellow and magenta distribution for CO₂ (2b) – which is describe as « slightly different » in the text but it is not really « slight » - at least give the exact number of the difference in the text.

b. I have difficulty to understand why there is no magenta rectangles below 450 ppbv on figure 2b (CH₄) while the difference between the blue and black rectangle on figure 2a are not so different.

9- Supplementary material

a. l. 49 : show the data or give some order of magnitude of the corrections which are done.

b. l. 141 :

: what do you mean by a preservation bias which is « uniform downcore ». It would be helpful to have the equations to understand/follow what you mean for all the section on the preservation bias.

c. You give different explanation of the d18O_{atm} not having an interglacial bias in the main text (diffusion effect) and in the supplement (precessional signal). Check coherency

Referee #2

(Remarks to the Author)

I thank the authors for their revised paper. Several issues have been explained more clearly and some interpretations have been toned down. The paper, with its improved dating compared to previous papers, and its use of direct ice core greenhouse gas measurements, remains a potential breakthrough that most likely warrants publication in Nature.

There remains one area of the paper where I feel the authors may not have fully understood what my concern was on the last round. This concerns the preservation bias. The authors go to considerable (perhaps too much) length to explain that the deep samples cover a very long period of time. I felt this went without saying, and myself I feel that Figure 3 does not deserve to be in the main text. However, this does not cover the issue about preservation bias, which has nothing to do with how much time each sample covers. The question is whether each sample averages across time in a consistent way. Looking at Fig 1a and b it's obvious, as stated in the paper, that there is a strong interglacial bias for the short snapshots. If one was to average all the samples from the short snapshots, the average would not be a time average of the EDC reference value (in fact Fig 2b, CO₂ panel, with a strong mismatch between the measured yellow and the magenta calculation, suggests that the interglacial bias is even stronger than the authors have used). However many of those short snapshots you take, if you simply average them without a correction, they will give you a value of around 250 ppm, not the EDC time average of around 220 ppm. This will still be true if you combine them by joining them together and thinning them, which is what a long snapshot is.

The authors suggest that the interglacial bias is due to the higher accumulation rates in the interglacial, ie interglacial time is way overrepresented in any length of ice. They also claim (partly from the data in Fig S4) that the interglacial bias persists throughout their record. But this cannot be true: when it comes to the long snapshots for the same time period, they do give roughly the EDC time average, implying that the interglacial ice is not overrepresented. Thinning and folding do not in themselves alter this. So what has happened – has some kind of differential thinning compensated so that there is (by chance) a time average in these long snapshots.

The reason why this is critical is because it begs the question whether the long snapshots below 800 kyr behave like the long snapshots or the short snapshots, and we simply don't know. At one end of options, the early Pleistocene snapshots may have an interglacial bias, they might have no bias, and they might even have a glacial bias. This is really not well-explained in the text, in fact not at all in the main text. One can partly understand this from Fig S5 but even there the issue is not made clear.

The authors really need to have another go at this and clarify it in the text. Do they have an idea why the interglacial bias is not seen in the long snapshots from 800-500 kyr, as it really should still be there unless another process, not mentioned, has taken place. Their preferred option is that all the long snapshots are unbiased recorders (ie without an interglacial bias), but I can't see why this should be. I am Ok with the options they end up with (although they don't cover the possibility of a glacial

bias in the early Pleistocene) but it is not well-explained, nor is it clear why they prefer the no-bias scenario.

A few more detailed comments:

Abstract, 117-19. As I say this seems like a reasonable conclusion but it is really not supported by what is in the text.

Line 23. <300 ppm doesn't seem to come from the text, at line 205, it says <280 ppm, and the uncorrected samples are all around 250 or less.

Line 138: "Good agreement between the sampled synthetic dataset and the distribution of Allan Hills samples for both 'short' (Fig. 2b) and 'long' (Fig. 2c)". The agreement for CO₂ in Fig 2b (short snapshots) is poor. Again I emphasise the key conundrum here is that the actual data (yellow histogram in CO₂ Figs 2b and 2c) are so different. No amount of thinning and averaging can turn the data in 2b to that in 2c for CO₂, so what has happened?

Line 162-4. A reader would never work out from reading the main text how you got to these numbers, or even what assumptions you have made. It needs a better explanation.

Line 212-3. This is probably for the editors but it is not OK to use this unpublished data here, so unless the MOT paper is also accepted this sentence should be removed and just the marine core evidence used at this point.

Line 239. I think you need to add a further caveat: the temperature could explain the CO₂ reduction but this is not only assuming sources and sinks in balance but also that ocean temperature change didn't affect other processes such as biological productivity, ocean circulation, etc.

S124-5. This seems circular to me: you are saying there is not a change in the bias because the temperature does what you are expecting, but this is also a result of the paper.

Line S174 and following better explains why you have different scenarios, but again if it is your preferred scenarios you need to explain why "long-exposure snapshots in near-basal ice that preserve climate signals that more closely represent the mean over a glacial cycle" when you have said that there is an interglacial bias throughout.

Fig S4: y-axis labels are wrong (0.1, 0.2, 0.3 each appear twice).

Version 3:

Reviewer comments:

Referee #1

(Remarks to the Author)

I read the answers to the comments as well as the new version. My comments were mostly taken into account, there are some inherent limitations linked to the nature of the ice cores but I feel that these results strongly deserve publication soon now.

(Remarks on code availability)

But I reviewed the data

Referee #2

(Remarks to the Author)

This third version of the paper is substantially revised from the last one, and does attempt to respond to all the points I (and the other reviewer) made on version 2. In doing so it takes a new tack: previously it noted the interglacial bias apparent in the short snapshots, but asserted that the long snapshots were some kind of average value of GHG concentrations. In the new version a substantial new section attempts to dismiss the short snapshots as somehow biased (either by being unrepresentative or through an addition of CO₂). This gives the authors a clear run to suggest that the long snapshots really are some kind of time average, and that this is likely to be unbiased through time. My feeling is that, had they decided on this route in version 1, much of the discussion of short snapshots might have gone into the supplement and we might have had a simpler and clearer paper, but I can see how we ended up where we are.

I think at this stage we have gone as far as we can in trying to unravel what the long snapshots really mean. I feel sure that everything I suggest the authors will already have thought of and discussed among themselves. The truth is that there are some mysteries in the data; we don't really understand the reason the short snapshots fall away from the long ones, and we don't really know what kind of average the long snapshots represent. However, the data deserve to be treated alongside the other major CO₂ datasets that are out there, and it's at least the simplest working hypothesis that whatever bias is present stays constant with time. Once we accept that this is the best that can be done at this stage the conclusions of the paper follow. I am therefore now recommending publication, although I do have a few suggestions about how the work is presented.

For me the critical point remains that the long snapshots are undoubtedly some kind of weighted average across the time

they represent. The physics of higher snow accumulation in interglacials makes a weighting towards interglacials a starting point, but ice dynamical differential thinning might lead to other weightings. I would therefore recommend that this is explicitly stated at key points, ie that the values are referred to as weighted averages that may not be simple time averages. This then also allows the authors to explicitly state that they have made an assumption that the weighting across climate stages is invariant. I will suggest in my detailed comments where this might be done.

The key conclusion of the paper is that there is only a small CO₂ trend across the Pleistocene (250 to 230 ppm). The authors acknowledge that preservation bias could account for up to 20 ppm. They don't really consider the fact that a trend of this order could also be induced by changes in cycle length and that shorter cycles before the MPT might also account for a part of the trend. Take a hypothetical situation where glacials were 200 ppm and interglacials 250 ppm throughout the Pleistocene. With 40 kyr cycles split 50:50 between glacial and interglacial and no preservation bias, the "average" would be 225 ppm, while with 100 kyr cycles split 80:20, the average would be 210 ppm, giving a 15 ppm trend that would be somewhat spurious. It would be useful to discuss this somewhere, even if using less simplistic calculations than I did.

Detailed comments:

Line 17: I suggest "likely represent some kind of weighted averages over glacial cycles. Assuming the weighting stays constant, our values permit direct reconstructions of atmospheric greenhouse gases..."

Line 129 "aye{Citation}rs" – check.

Line 119-155. While I understand this discussion it is a lot of text in the main paper, and quite peripheral to the main discussion as your conclusion from it is to ignore whatever the short snapshots are telling you. I don't object to it being here but wonder whether there is material that would be more important in the main text than this?

Line 131-132: Your suggestion about the use of 38Ar doesn't really work, because you explicitly believe the short snapshots are mainly interglacial while the long snapshots are means, and yet Fig S9 shows an unchanged 38/36Ar value between short and long snapshots of the same age. This implies that this measure is not diagnostic of climate state. Unless you can explain this observation you need to downweight this evidence.

Line 132: you mean Fig S9 not Fig S8.

Line 158: I suggest a couple of additions (which I have placed in brackets): "where the oldest ice preserves CO₂, CH₄, and $\delta^{18}\text{O}_{\text{atm}}$ averaged (with an undetermined weighting between glacial and interglacial) over one or more glacial cycles. Although this interpretation of the Allan Hills ice core records limits our ability to comment on the amplitude of glacial-interglacial variability in these properties, we believe they provide robust records of (weighted-)mean atmospheric composition for CO₂ and CH₄ over long intervals..."

Line 166: "This bias does not appear to change with sample age". While I agree this is your assumption I don't see on what basis you are asserting it to be a fact. Clearly we have no way of knowing whether interglacials are overrepresented in some samples and not others, and at some ages and not others. What your calculations show is that there can be a bias of up to 20 ppm, and we have no way of knowing which samples this applies to. Your reasonable assumption is that it stays the same at all ages, but have you really done anything that allows you to assert that it appears to not change with time?

~line 185. Would you like to add a sentence reflecting my comment about the proportion of time spent in glacial and interglacial states.? I think I can accept if you prefer not to add this complication.

Line 188. I wonder if it would be worth adding that your values fall within the range of 11B data, but for them to be compatible with 11B-based reconstructions, your samples would have to have a massive bias towards glacial ice, which seems improbable.

Line 247: I suggest "likely reflecting atmospheric compositions averaged with some kind of climate weighting over glacial cycles". (Since in the next sentence you say they have a bias they are clearly not a mean in the sense most readers would envisage it). And then "a reasonable assumption is that this bias does not change with sample age".

Ref 13 is incomplete but i believe is now in press.

Supplementary. The discussion of preservation bias that leads to Figs S7, S8 and S5 is very hard to follow. I just about grasped it but it is not straightforward. Please take one last look to see if you can explain it more simply: a start might be to reorder the figures so that we can see the scenarios you are calculating (S7, S8) from before we see the result in S5!

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

In responding to reviewer comments we have made substantial changes to the content and organization of the main and supplementary documents. Several figures have changed. Where a figure has moved, we state “former Fig. X” in our responses to indicate the figure location in the original manuscript. Below is a table listing a summary of the changes:

Description of change	New manuscript	Former manuscript
Added section discussing various interpretations of signal preservation	-New section in supplement from line 98-187 (also pasted below in responses to reviewers) -Summary of findings from supplement can be found in main text lines 126-163	-Did not exist in former manuscript
Added and changed figures associated with new section about signal preservation	-New figures associated with this section are main text figures 2 and 3, and supplementary figures 2, 4, and 5.	-Supplementary figure 2 has been revised and moved to the main text (Figure 2a). A second experiment distinguishing between short and long snapshots is included (Figure 2b-c) - Supplementary figure 3 is now part of figure in the main text (Figure 2d-e)
Added uncertainty associated with record to figure 1	-Figure 1: linear regression through CO ₂ and CH ₄ removed, additional annotation with results for long-term CO ₂ and CH ₄ change depending on interpretation	
Added map of Allan Hills and core sites	-Supplementary figure 1	-did not exist in former manuscript
Moved section about redating ALHIC1503	-Supplement (starting line 73)	-Main text (starting line 138)

Referee #1 (Remarks to the Author):

This manuscript builds on data obtained from new ice cores drilled in the Allan Hills region enabling to get snapshots of old atmosphere and hence to infer the evolution of CO₂ and CH₄ in the atmosphere over the last 3 millions years. The main finding is that the CH₄ level did not significantly change over the last 2.9 millions years and that CO₂ level only slightly decreased by approximately 25 ppm.

This study complement the study of Yan et al. (2019) by adding more core and significantly more dated ice level. The conclusion is also a bit different. While Yan et al. (2019) commented on the reduced amplitude of the amplitudes of the glacial – interglacial cycles for both climate and greenhouse gases concentration, the present manuscript only comment on the « mean » level (which may actually be biased toward interglacial values).

This study has a broad enough implication to be of high interest for publication in Nature. Still, there are many points requiring further investigations or discussions in order to prevent misinterpretation.

We appreciate the feedback provided by this reviewer and hope we have made satisfactory changes that address their concerns.

1- The authors did some efforts in the companion paper to explain the different drilling also including a map in the supplementary information. Still, it could be nice to have it here also with a clear indication of what has been measured in each cores used here for the synthesis (ALHIC 1502, 1503, 1901, 1902, 1903), especially the age – depth relationship for each core. This could also help to understand why the conclusion from this manuscript is different from the conclusion of the paper of Yan et al., 2019. The differences in interpretation (especially about the amplitude of the glacial – interglacial cycles) should also be explained in the manuscript to help following the research on these ice cores network.

We recognize that without being able to see one of the papers originally submitted as a third companion paper, some fundamental details on the setting were lost. A detailed map of each ice core site is provided in Shackleton et al. 2025, currently in review at PNAS and publicly archived on ESS Open Archive (DOI: 10.22541/essoar.173897232.24923969/v1). We have added a map of the relevant cores to this study as a figure in our supplementary document as well.

2- I am missing a reconstruction of the age for the different ice level from the measurements of the d40/38 and d38/36. The data are never shown nor the limitations associated with the method (and discussed in other previous publications such as Yau et al.,). However, when looking at the data provided as a table, we see that ages of several millions years associated with strange d38/36 and d40/36 data were removed but there is no explanation on why / how there was a selection in the data shown here. Recent communications have shown that ^{81}Kr dating can be performed on samples of 1-1.5 kg. It would be very helpful knowing the limitations with the $^{40}\text{Ar}/\text{atm}$ dating to provide a few ^{81}Kr dates.

As mentioned in the response to the previous question, we recognize reviewers could not see the paper reporting ages for the samples discussed in here and we direct current and future reviewers to the ESS Open Archive where Shackleton et al. 2025 has been uploaded (DOI: 10.22541/essoar.173897232.24923969/v1). We are aware of advances in ^{81}Kr dating but unfortunately at this time US State Department rules do not allow us to ship samples to the lab in China that would be able to do this work. We hope to resolve this problem in the future but it is not feasible to make these measurements now.

3- When looking at the d38/36 data in the supplementary, there is no link between these data and the d18Oice. This is a bit surprising since there should be a link because of gravitational fractionation. Have you investigated it ? Any explanation ? Also the d38/36 data for the top part are quite high compared to the rest. Again any explanation for these data ? May it influence the $^{40}\text{Ar}/\text{atm}$ dating ? The d38/36 data should be discussed a bit in the supplement at least (either here or in the companion paper).

d38/36Ar data were not in our original manuscript, though they are now plotted in a supplementary figure to discuss changes in climate-dependent preservation bias down core (more details can be found in response to question 4). Otherwise, a discussion of d38/36Ar can be found in detail in Shackleton et al. in Review (publicly archived on ESS Open Archive (DOI: 10.22541/essoar.173897232.24923969/v1).

The d38/36Ar represents the gravitational fractionation in the firn column, and like the more widely-used $d^{15}\text{N}$, is a measure for the thickness of the firn diffusive zone (Sowers et al., 1992). The lowest d38/36Ar values cluster around 0.15‰, or a 15m thick diffusive zone. Firn density measurements of firn cores from the Allan Hills accumulation zone suggest a present-day air lock-in depth of around 20 m. The difference between these two could be explained by vigorous air ventilation of the upper 5 m of the firn column, in the so-called convective zone.

The thickness of the firn layer is controlled chiefly by two climatic parameters – temperature (T) and snow accumulation rate (A). These have an opposite effect on firn thickness, with increasing T and A resulting in a thinner and thicker firn, respectively. As T and A tend to co-vary through time, changes in firn thickness are relatively minor on glacial-interglacial timescales as the T and A impacts cancel out to first order. For example, comparing the last glacial maximum (LGM) to the preindustrial firn thickness, we find a 25% increase at the WAIS Divide ice core, a 12% decrease at the Dome Fuji core, and no significant change at Talos Dome (Buizert et al., 2021).

The relationship between $d^{18}\text{O}_{\text{ice}}$ and $d^{38/36}\text{Ar}$ at Allan Hills is plotted in figure R1 below, separated into the 1903 core (purple, this core resolves termination T2 in stratigraphic order with minimal smoothing), 1902 core (orange), and the short (green) and long (blue) snapshots for the 1901 cores. Note that no data from ALHIC1902 or ALHIC1903 was used in this paper but can be found in Shackleton et al. 2025. For both the 1903 and short snapshots, we find a weakly positive correlation, that suggests a dominance of the impact of A over T on glacial-interglacial cycles as is also observed in the Antarctic interior (for example, Dome C, Dome F, Vostok, EDML, South Pole), and reported previously for Allan Hills (Yan et al., 2021). This behavior falls well within the range of observed trends as noted above. The short snapshots show a reduced range of values in both $d^{18}\text{O}_{\text{ice}}$ and $d^{38/36}\text{Ar}$ relative to 1903, consistent with the idea that these samples have experienced some smoothing, but still maintain some glacial-interglacial variance. Note that 1903 is isotopically colder in $d^{18}\text{O}_{\text{ice}}$ than 1901 and 1902, which is consistent with the thicker firn thickness in 1903.

For the long snapshots, the glacial interglacial cycles are removed via smoothing and we no longer observe the positive correlation between $d^{18}\text{O}_{\text{ice}}$ and $d^{38/36}\text{Ar}$ for the Pleistocene ice (roughly $d^{18}\text{O}_{\text{ice}} < -35.5\text{‰}$). Instead we find a weakly negative correlation, that would imply that on longer timescales the temperature effect dominates. For the Pliocene ice (roughly $d^{18}\text{O}_{\text{ice}} > -35.5\text{‰}$), we find a very strong positive correlation. At face value, this would imply a much thicker firn column under Pliocene warmth, most easily explained via strongly enhanced accumulation rates in the early Pliocene in a way that scales non-linearly with site temperature. These oldest samples from ALHIC1902, however, have enriched $d\text{Kr}/\text{Ar}$ ratios consistent with argon loss that may enrich $d^{38/36}\text{Ar}$ artefactually (Buizert et al., 2023). More work is ongoing to understand Pliocene firn thickness. Any argon isotopic fractionation during gas loss is expected to scale in a mass-dependent

way, and therefore does not to first order impact the Ar dating, in which such effects are corrected for. Please refer to Shackleton et al. in review for further detail.

The reviewer notes that the $d^{38/36}\text{Ar}$ data appear higher in the shallower ice. We assume that they assessed this by plotting the depth column vs. the $d^{38/36}\text{Ar}$ column in the Excel sheet, in which case all the cores are combined and the interglacial (MIS5e) ALHIC1903 data (which are indeed higher in $d^{38/36}\text{Ar}$) appear as the shallowest data. The shallowest ALHIC1901 $d^{38/36}\text{Ar}$ data point is also quite high, but it is isotopically warm in $d^{18}\text{O}_{\text{ice}}$, suggesting it likewise to be an interglacial sample.

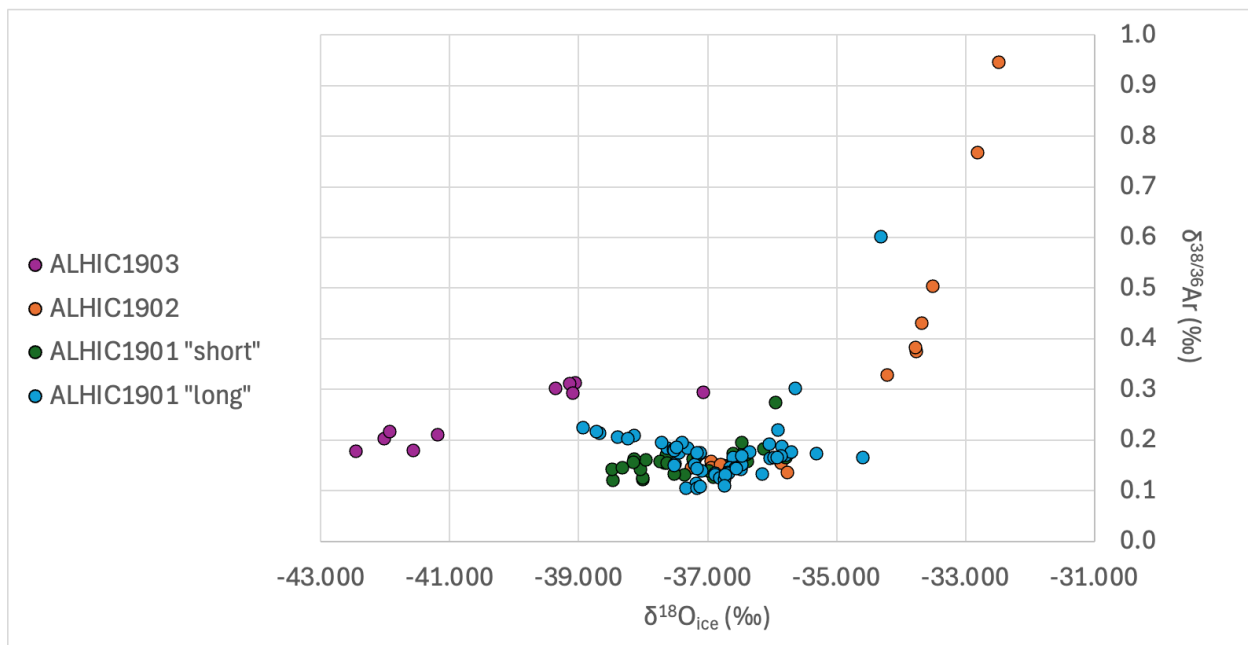


Figure R1: the relationship between $d^{18}\text{O}_{\text{ice}}$ and $d^{38/36}\text{Ar}$ in Allan Hills cores. Broadly, Pliocene and Pleistocene samples have $d^{18}\text{O}_{\text{ice}}$ values that are above and below around -35.5 ‰, respectively.

Buizert, C., Fudge, T.J., Roberts, W.H.G., Steig Eric, J., Sherriff-Tadano, S., Ritz, C., Lefebvre, E., Edwards, J., Kawamura, K., Oyabu, I., Motoyama, H., Kahle, E.C., Jones, T.R., Abe-Ouchi, A., Obase, T., Martin, C., Corr, H.F.J., Severinghaus Jeffrey, P., Beaudette, R., Epifanio, J.A., Brook, E.J., Martin, K., Chappellaz, J., Aoki, S., Nakazawa, T., Sowers, T.A., Alley, R.B., Ahn, J., Sigl, M., Dunbar, N.W., Svensson, A., Fegyveresi, J.M., He, C., Liu, Z., Zhu, J., Otto-Bliesner, B., Lipenkov, V.Y., Kageyama, M., Schwander, J., 2021. Antarctic surface temperature and elevation during the Last Glacial Maximum. *Science* 372, 1097-1101.

Buizert, C., Shackleton, S., Severinghaus, J.P., Roberts, W.H.G., Seltzer, A., Bereiter, B., Kawamura, K., Baggenstos, D., Orsi, A.J., Oyabu, I., Birner, B., Morgan, J.D., Brook, E.J., Etheridge, D.M., Thornton, D., Bertler, N., Pyne, R.L., Mulvaney, R., Mosley-Thompson, E., Neff, P.D., Petrenko, V.V., 2023. The new Kr-86 excess ice core proxy for synoptic activity: West Antarctic storminess possibly linked to Intertropical Convergence Zone (ITCZ) movement through the last deglaciation. *Clim. Past* 19, 579-606.

Sowers, T., Bender, M., Raynaud, D., Korotkevich, Y.S., 1992. $\delta^{15}\text{N}$ of N_2 in air trapped in polar ice: A tracer of gas transport in the firn and a possible constraint on ice age-gas age differences. *J. Geophys. Res.* 97, 15683-15697.

Yan, Y., Spaulding, N.E., Bender, M.L., Brook, E.J., Higgins, J.A., Kurbatov, A.V., Mayewski, P.A., 2021. Enhanced moisture delivery into Victoria Land, East Antarctica, during the early Last Interglacial: implications for West Antarctic Ice Sheet stability. *Clim. Past* 17, 1841-1855.

4- I am surprised by the assumption that CO_2 and CH_4 « values agree well with depth-weighted averages of these gases from the same period in EDC (sup figure S3) ». Indeed, on his figure, there is as important proportion of the red stars falling on high level of CO_2 and mid-level of CH_4 (such CH_4 - CO_2 characteristics are rare in the EDC records). This is surprising and should be explained because it is also linked to the main conclusion of the manuscript stating that CO_2 levels are a bit higher than pre-industrial levels for old period while CH_4 levels are similar to pre-industrial level.

We agree that this is an important section of the paper that can be emphasized more. To answer the first part of this question, our assumption that CO_2 and CH_4 measurements plot closer to the EDC average comes from the observations from deep samples only (>141 m), or what we deemed “long-exposure snapshots”. In our new figure (Main Fig. 2d-e, formerly supplementary Fig. 3), these are represented by the blue circles. The red circles in that figure are from shallower depths where we observe more variance in both CO_2 and CH_4 . These samples we designate as “short-exposure snapshots”. In the new version of the paper, we have added additional justification for this depth-dependent change in variance by calculating the age density, or the amount of time represented by a unit of depth (yr/cm). Age density is shown in our new figure (main Fig. 3). A jump in age density by roughly one order of magnitude occurs at approximately 141 m.

The second part of this question is also important to discuss - in the shallower samples (“short snapshots”) we tend to measure higher CO_2 for a given CH_4 measurement, compared to EDC. We have examined the possibility that there is a small amount of CO_2 production in these samples, enriching the CO_2 measurement from its atmospheric concentration. However, the isotopic signatures of CO_2 from these samples are within the limited range we observe for

atmospheric CO₂ in the traditional ice core record (-6‰ to -7.1‰). Alternatively, through our smoothing experiments, we show that it is more likely to preserve high CO₂ concentrations than CH₄ concentrations when these signals have been heavily attenuated (main Fig. 2b, formerly supplementary figure 2). This is because, in general, CO₂ levels remain higher for longer than CH₄ levels during interglacial periods. This effect is illustrated in the new main figure 2 and the analysis of its impact is described in the supplement.

We have moved a more detailed discussion of the comparison between ALHIC1901 and EDC gases into the main text and added a new figure to the main text with the results of the various smoothing experiments described above.

Lastly, though not directly addressing the question here, in response to a comment from Reviewer 2 we have made some edits to the framework used to discuss the extent of thinning/physical mixing and preservation bias in our record, incorporating more uncertainty, and these changes may also be relevant to the questions about smoothing and averaging. Though our comparisons of CO₂ and CH₄ concentrations downcore compared to EDC have led us to infer that the deeper samples represent long-term averages that have smoothed out orbital-scale variations, we recognize that there might be alternative interpretations. In the supplement, we now discuss a range of scenarios associated with the 1) extent of a climate-dependent preservation bias and 2) the extent of physical mixing and/or layer thinning. We present a most likely scenario, but include the range of possibilities for CO₂ histories in Figure 1 (as annotations). By incorporating these other scenarios, we are bounding our record with uncertainty that we believe will satisfy the concerns discussed here. It is important to note that CH₄ concentrations remain quite stable across the Pleistocene, no matter which scenario is used. The new section of the supplement has been pasted below, and can be found starting on line 89 of the new supplementary document.

"Preservation of atmospheric signals in ALHIC1901

ALHIC1901 measurements of CO₂, CH₄, and $\delta^{18}\text{O}_{\text{atm}}$ for samples with $^{40}\text{Ar}_{\text{atm}}$ ages between 400-800 ka recover $42 \pm 1\%$, $38 \pm 1\%$ and $23 \pm 1\%$ of the EDC range for that time period, respectively. These fractions were derived by first interpolating each EDC record¹⁸⁻²⁰ to a 50 year even spacing. We then created 10,000 Monte Carlo simulations of each EDC record, incorporating published analytical uncertainty¹⁸⁻²⁰. We report the fraction that the range in values for each parameter in ALHIC1901 represents relative to the average range of the Monte Carlo simulations. These ranges were calculated using all ALHIC1901 measurements from 400-800 ka, regardless of depth.

When considering how to extrapolate the extent of preservation bias and signal attenuation (due to physical mixing and/or thinning) to measurements older than EDC, we separated the two processes and assessed their individual impacts on the gas records:

1. Preservation Bias

For samples older than the continuous ice core record we examine two endmembers: 1) uniform preservation bias through time, or 2) greater preservation bias in the late Pleistocene than during the late Pliocene/early Pleistocene. The latter scenario would arise if the glacial-interglacial accumulation contrast was greater in the late Pleistocene, a reasonable assumption considering that early Pleistocene and late Pliocene glacial minima were likely less extreme. Under this assumption, values measured at 2.9-2.7 Ma would represent unbiased long-term atmospheric averages. Comparing the mean concentrations for CO₂ and CH₄ (250 ppm and 507 ppb, respectively) to the atmospheric averages during the late Pleistocene (220 ppm and 503 ppb, respectively) indicates a CO₂ decline of 30 ppm and no significant change in CH₄.

Evidence for a change in preservation bias may be observed in the $\delta^{38/36}\text{Ar}$ ratio, which records changes in the height of the diffusive zone within the firn column via gravitational enrichment. Most if not all polar ice cores show a glacial-interglacial difference in firn thickness. The value of $\delta^{38/36}\text{Ar}$ is relatively constant throughout our whole record, indicating little change in the gradient between glacial and interglacial layer thickness (Supplementary Fig. 4). The stability in $\delta^{38/36}\text{Ar}$ would be consistent with a preservation bias that is constant throughout the record, though it does not provide definitive proof thereof.

Lastly, though a preservation bias toward glacial periods is not observed in the younger part of our record, we acknowledge that it is possible due to differing ice rheology leading to a change in ice flow between interglacial and glacial layers (Durand et al. 2007). Evidence against this hypothesis comes from the expected cooling trend across the Pleistocene observed in both the water isotope record (Shackleton et al. in review) and mean ocean temperature record (Shackleton et al. in review) from the same ice core. In addition, we would expect to see a difference in $\delta^{38/36}\text{Ar}$ if there was a switch in preservation bias from glacial layers in the early Pleistocene toward interglacial layers in the late Pleistocene.

2. Thinning/physical mixing

To examine the impact of assumptions about thinning and/or physical mixing, we consider two endmembers: 1) either the extent of physical mixing and/or thinning is depth dependent, or 2) the whole core has experienced the same amount of signal smoothing. First, we examine whether the extent of smoothing changes with depth. Plotting the gradient in age vs depth ("age density") against depth shows a change around 140 m, below which the amount of time represented by one centimeter of ice increases by an order of magnitude (Fig. 3 inset). A change in variance of CO₂ vs CH₄ also occurs at approximately this depth (141 m) (Fig. 2d). Ice younger than 800 ka but stratigraphically below this depth threshold records much less variability in all climate properties. The transition coincides with both the appearance of ice older than 1 Ma and evidence for contamination of trapped CO₂ by respired organic carbon (~141 m depth). We can replicate this change in signal attenuation by employing different smoothing above and below 141 m (Gaussian windows differing by an order of magnitude; Fig. 2b-c). Because we believe the preservation bias is uniform downcore, we again used the synthetic EDC record weighted to S27 layer thickness as described in the main text. We generally are able to reproduce the observed distributions above and below 141 m ("short" vs "long" snapshots). $\delta^{18}\text{O}_{\text{atm}}$ does not show an interglacial bias in the same way as CO₂ and CH₄ which is similar to what we see in the Allan Hills record (Fig. 2). This is likely because the 800 ka $\delta^{18}\text{O}_{\text{atm}}$ record resembles precessional variations²⁰⁻²² whereas CO₂ and CH₄ are dominated by 100 kyr periodicity in the late Pleistocene.

Results show a slight disagreement between the model distribution of CO₂ and the ALHIC1901 measurements above 141 m (Fig. 2b). However, considering the much shorter amount of time likely represented by the shallower samples, a much larger dataset than we have available would likely be required to be certain of this result. Furthermore, if the stratigraphy above 141 m is dipping, we would expect to measure less variability across depth. Thus, it is possible that a higher resolution dataset in this shallow section would exhibit a more representative distribution of EDC. Alternatively, the larger variance in CO₂ relative to CH₄ and $\delta^{18}\text{O}_{\text{atm}}$ could partially be explained by small deviations from true atmospheric CO₂ values. A number of processes have been observed to alter the CO₂ concentration in ice cores, such as the presence of melt layers, carbonate or silicate weathering, oxidation of organic carbon, and post-coring effects^{23–27}. Apart from silicate weathering and carbonate precipitation, most known processes increase the CO₂ concentration in bubbles leading to an overestimate of atmospheric CO₂. While in some cases CO₂ can be highly enriched (for example, due to oxidation of organic matter near the bedrock interface), in other cases alteration is subtle, with differences between ice cores reported in the 5–20 ppm range^{8,19,28}.

The change in age density in ALHIC1901 indicates that samples below 141 m (which include all samples older than 1 Ma) may have experienced roughly 10 times more signal attenuation than the samples above 141 m. With each sample below 141 m representing more time, a smaller sample size may achieve a more representative distribution, which may explain why we see good agreement between the experimental distributions and the deep Allan Hill samples (Supplementary Fig. 2c). Thus, there appear to be at least two types of paleoclimate snapshots in our record: 1) short-exposure snapshots in shallow ice that preserve some of the true glacial-interglacial variability, and 2) long-exposure snapshots in near-basal ice that preserve climate signals that more closely represent the mean over a glacial cycle. Examining just the long-exposure snapshot record reveals a 20 ppm decline in CO₂ from 2.9–1.2 Ma followed by stable CO₂ concentrations through the MPT and into the late Pleistocene, and no change in CH₄ (Fig. 1). Considering the sample number gets smaller with increasing age in the current dataset (late Pliocene samples: n = 2), we also analyzed the long-exposure snapshot record by binning samples in time and conducting statistical t-tests assuming unequal variance. The mean CO₂ and CH₄ values were significantly different (p-value < 0.05) between the pre-MPT (>1.5 Ma, n=10) and post-MPT (<0.8 Ma, n = 16) samples (Supplementary Fig. 5), suggesting a drop in CO₂ and CH₄ of 10 ppm and 10 ppb between these two binned time periods, respectively.

As an alternative approach, we can assume the time represented by each sample is the same for the entire record. Conducting a t-test between early Pleistocene (older than 1.5 Ma) and late Pleistocene (400–800 ka) samples shows lower CO₂ in the older samples and an insignificant difference in mean CH₄ (Supplementary Fig. 5). Under this endmember scenario, CH₄ shows no change from 2.9–0.8 Ma while CO₂ exhibits a ~7 ppm increase.”

Figure 1 in the main text has been edited to include these scenarios as written annotations, providing uncertainty bounds on the secular decline observed across the Pleistocene. The main text has been edited to include this uncertainty as well, stating the long-term change in CO₂ observed is from +7 to -30 ppm from 2.9 to 0.8 Ma. While we wholeheartedly agree that incorporating uncertainty into our record is important, our data do not support that CO₂ was as high in the early Pleistocene as indicated by average estimates from boron isotope proxy reconstructions.

5- The study related to figure S2 should be improved for two reasons. First, it seems that the authors assumed the same diffusion for CO₂, CH₄, d18O_{atm} while the diffusivities and solubilities are probably different (see Bereiter et al., 2009). This should be calculated / discussed. Second, I am not really sure how you want to extract interglacial values for d18O_{atm}. The beginning of the interglacial periods (according to CO₂ – CH₄) show high d18O_{atm} (1 permil or so) and then d18O_{atm} decreases. The initial distribution of d18O_{at} is also different from the initial distributions of CO₂ and CH₄ with very few high d18O_{atm} values – this could also affect the comparison that you are discussing.

We recognize that the figure representing these experiments was not the most intuitive.

To answer part 1 of the question, ice from this core has very likely undergone extreme amounts of smoothing, due to ice flow histories leading to layer thinning and potential physical mixing of layers. An impact from molecular diffusion is also a possibility, however work done by Yan et al. 2019 calculating the diffusion timescale of molecular oxygen and nitrogen in ice at Allan Hills temperatures suggests diffusion on the submeter length scale would take 10⁷ years. Considering the smaller molecular diameter of O₂ and N₂ compared to CO₂ and CH₄ (Ikeda Fukuzawa et al, 2001, Bereiter et al. 2014), we assume molecular diffusion is not the dominant process producing the observed smoothed signals. Thus, in the experiment that produced Fig. S2, we uniformly smooth each synthetic EDC record assuming the physical mechanism(s) responsible for smoothing would have the same effect on each gas parameter.

Part 2 of the question asks about extracting interglacial values from d18O_{atm}. In our smoothing experiment, d18O_{atm} does not show an interglacial bias in the same way as CO₂ and CH₄ which is similar to what we see in the Allan Hills record (Fig. 2). This is likely because the 800 ka $\delta^{18}\text{O}_{\text{atm}}$ record resembles precessional variations whereas CO₂ and CH₄ are dominated by 100 kyr periodicity in the late Pleistocene.

We have remade this figure to make the distributions of each gas parameter under the various smoothing scenarios clearer. This is now a figure in the main text (main figure 2, former supplementary figure 2), and a brief description of this experiment has been moved from the supplement into the main text. We also add discussion about why d18O_{atm} does not show interglacial bias.

6- Figure 1 – I am unable to see the « open blue circles » referred to in the caption

This was an error, the caption was not updated from an earlier version. It has been changed.

7- I am surprised by the mention of negative Ar/N₂ ratio ? This is not possible ? You mean significantly low dAr/N₂ ? Could you provide some more explanations in the text by giving values. Again, there is a lack in showing the elemental and isotopic values for argon.

Thank you for this comment - anomalous noble gas ratios are seen in the bottom 0.5 m of ALHIC1901 and described in detail in the companion paper on mean ocean temperature (Shackleton et al. 2025). The statement in our supplement was incorrect because it was missing a delta symbol. This section has been changed to the following, starting on line 209 in the supplement: “Using noble gas ratios as a proxy for melt processes, we can rule out dissolution as an explanation for elevated CO₂ concentrations, as this would lead to highly enriched $\delta\text{Xe/Kr}$ and $\delta\text{Kr/Ar}$ relative to our atmospheric standard (Orsi et al, 2015), which is not observed. Evidence of gas alteration in the deepest sample (⁴⁰Ar_{atm} age of 4 Ma), indicated by highly enriched $\delta\text{Kr/Ar}$ and depleted $\delta\text{Xe/Kr}$ (Shackleton et al. 2025), opens up the possibility of secondary processes altering the CO₂ concentration, and we therefore exclude this deepest sample from our record.”

8- The correction for CO₂ levels using d13Catm is simple but I agree that the authors can not really do better. It is important to show on the main figure 1 that you have different types of data. Also this biological effect can probably also affect CH₄ and d18Oatm and was also taken into account ? Is it possible to discuss this a bit or are the samples removed from the analysis ?

All CH₄ and d18Oatm measurements were excluded from figures and interpretation where a secondary source of CO₂ was identified. This point was originally made in the supplement and has been moved to the main text, as well as illustrated in Fig. 3 where all the open circles are samples containing non-atmospheric CO₂. All data can be found in the supplementary dataset.

9- There are some strange values in the ALHIC1503 series display on supplementary figure 1 : very low dO₂/N₂ data, very low d15N – what was done with these data ? I imagine that they are not further used here because the ice is not dated but it would be nice to comment since you are showing them.

This figure has been updated to exclude ALHIC1503 data that were not able to be interpolated because they were outside the depth range measured in ALHIC1901. The only data included on the figure are ALHIC1503 samples that were redacted

using our new chronology. This figure is now Supplementary Fig. 6, former supplementary fig. 1. The values mentioned in this question as strange (low dO_2/N_2 and low $d^{15}N$) are from a shallower section of the ALHIC1503 core (previously known as the BIT58 core (Higgins et al. 2015)) where we do not have analogous ice in ALHIC1901. These data were measured with a different procedure than the deeper samples and the strange data are likely a result of gas loss and/or samples containing visible fractures (only fracture-free ice was sampled in the deeper section of the core). Details on these data can be found in Yan et al. 2021.

Referee #2 (Remarks to the Author):

This paper presents a new dataset of CO_2 and CH_4 values from blue ice cores from Allan Hills, Antarctica. The novelty is that the data extend back to 3 Ma, well beyond the conventional ice core time period. Compared to an earlier paper (Yan et al 2019) that extended 2 Ma in the past, there is a greatly improved dating, allowing an attempt to define trends with time, and the ability to compare with datasets of CO_2 from marine proxies and from models.

If it can be taken at face value, the results shown here (notably no change in CO_2 over the MPT, and a slow decline between 3.0 and 1.2 Ma) are very exciting and important. Given that they come from the more direct measurement of air in ice, they would effectively supersede and arbitrate between earlier compilations from different marine methods. However there are some issues that are treated in the supplement but glossed over in the text that might cause us to question what we are actually seeing here and whether it is an objectively superior datasets to the others. I don't think these preclude publication of the data, but they do require the authors to be rather clearer about the different options the dataset might represent and they make it less of a clear breakthrough than might be implied on a straight reading of Figure 1.

We have rewritten the main text (line 93-172) and supplement (line 98-187) to include a more detailed discussion of the various ways to interpret gas records from Allan Hills ice. The new section from the supplement is pasted below.

The issue is what kind of an average each datapoint represents and whether that average has a meaning that is constant as we go back in time. This is covered cursorily in lines 113-136 of the main paper, and more extensively in lines 90-164 of the supplement. The issue arises from the fact that only long-term averages can be derived in the older samples, and we would ideally like them to be reliable time averages of periods around the $40Ar$ age estimate. However such mixed samples can suffer biases from two causes. Firstly there is more snowfall in warm than cold periods, so we expect

averages that cover glacial cycles to be biased towards interglacial values. Secondly there are clearly flow disturbances in the ice, and experience in Greenland and Antarctica show that this can cause distortion of the age scale particularly around the interface between glacial and interglacial ice (due to their different rheology). This second issue might in theory lead to biases either towards glacial or interglacial values.

As mentioned above, the new version of our manuscript includes an expanded section on gas record interpretation, including a detailed discussion of the possible ways our measurements are biased. We use the overlapping data with EPICA Dome C records to assess possible biases and the implications these interpretations have on what the late Pliocene/early Pleistocene CO₂ and CH₄ measurements represent. As the reviewer notes, we find our record likely contains an interglacial bias due to the difference in amount of interglacial/glacial snowfall. We discuss two endmember scenarios where this preservation bias is either uniform downcore or age-dependent. This section is pasted below and figure 1 now includes written annotations to represent the uncertainty associated with these interpretations.

The only direct evidence the authors have about this issue is in the ice they refer to as “short snapshots”, which I believe is the ice above 141 m in ALHIC1901 (red in Fig S3B). There they observe a clear interglacial bias (using EDC as reference), which they ascribe to an extreme accumulation rate contrast between interglacials and glacials at Allan Hills. This explanation (also partly justified in Fig S2) seems reasonable. But then when they reach the deeper ice (146-148 m and ~151 m) of similar age, they find values (in a range of parameters) closer to the average at EDC. This is curious, and the authors use it to justify an assertion that all the values below 141 m are similar to time averages. This is really quite a strong assumption on very weak evidence and poses a number of challenges:

- a) If there is a strong interglacial bias in accumulation rate it must surely apply to the area of origin of all the Allan Hills ice, so why is the bias seen above 141 m not present below 141 m? The authors talk about a weaker accumulation contrast in older ice (line 157-160 of supplement), but this doesn’t explain the apparent reduced bias in ice younger than 800 ka below 141 m. (And even a 3:1 contrast as opposed to a 10:1 contrast in accumulation would impose a significant bias.)
- b) The most obvious answer to (a) is that the deeper ice also suffers a flow disturbance that has prioritised glacial ice, and therefore partly restored the balance between glacial and interglacial material.

c) If a and b are both in play then what is the justification for assuming that the older ice (142-146 m, 148-151 m, and below 152 m) has the same bias as the ice between 146 and 148 m?

d) A final issue, but perhaps less fundamental, is that the ice before the MPT (with what in benthic isotopes looks like symmetric glacial/interglacial cycles) imposes a fundamentally different averaging compared to the sawtooth shaped cycles of the post-MPT, which if there was no accumulation bias would have an average tending towards glacial values.

Thank you for these thoughts. First, we find that, most likely, the climate-dependent preservation bias that we see in the shallower section (“short” snapshots) is unchanged downcore. Evidence for this comes from the stable values of $\delta^{38}/^{36}\text{Ar}$, which records changes in the height of the diffusive zone within the firn column. The relatively constant values throughout the record indicate that the gradient between glacial and interglacial layer thickness is stable with time in this core. This is discussed in the new section below and can be found in Supplementary Figure 4. The alternative scenario is also discussed, as the reviewer notes in (d), where less extreme glacial minima in the late Pliocene/early Pleistocene could have resulted in a smaller preservation bias compared to the late Pleistocene. This is included in our uncertainty of the secular change in CO₂ across our record and written as an annotation in Figure 1.

Though preservation bias is likely uniform downcore, the extent of signal attenuation from physical mixing and layer thinning is likely an order of magnitude larger. Evidence for this comes from a calculation of age density, or the change in time over change in depth (dt/dz) based on the ALHIC1901 chronology. At ~141 m, age density increases by approximately an order of magnitude (new figure 3 in Main text).

We have now conducted a more robust experiment to consider changes in signal attenuation downcore (pasted below as well). Figure 2 in the main text now includes results from this experiment (and described in the main text starting on line 126 and the supplement starting on line 128). Figure 2b and c show the distribution of CH₄, CO₂ and d18O_{atm} from smoothed versions of EDC that differ by an order of magnitude (smoothing of 2c is 10 x smoothing of 2b). The agreement in 2c indicates that signal attenuation is the dominant mechanism impacting the “long” snapshots, even though a slight preservation bias may still exist.

The result of these issues is that it raises the fundamental question whether the averages have the same meaning as we progress back in time, which then raises

questions about the trends. At one extreme, one could assume that the accumulation rate bias seen in the short snapshots applies to older ice, and that the long snapshots of younger ice below 141 m are anomalous. This would lead to a very uncomfortable conclusion about interglacial concentrations but cannot be entirely discounted. At the other extreme one can assume that somehow (though the paper doesn't explain how) the accumulation rate bias is completely lost in the long snapshots younger than 800 ka, and that this same correction of the bias applies throughout the ice below 141 m (and its equivalent in ALHIC1503). This also seems implausible to me, but is the interpretation the authors have adopted. It leads to the also rather uncomfortable conclusion that the higher boron values, including some around 1.2 Ma (Fig S6, pink and purple averages), are wrong.

More realistically one could propose that in the older ice the values are closer to an average (due to a smaller accumulation contrast, and likely a weaker rheology contrast), but could contain a bias towards either glacial or interglacial values, imposing an uncertainty of probably tens of ppm on the numbers. Since this bias might vary systematically with depth and age, this will make it hard to define the trend from 3 Ma, or confirm the lack of trend across the MPT. Possibly that weakening of the message makes this a paper that warrants publication somewhere other than Nature, but that would be an editorial decision.

My recommendation therefore is that the authors should discuss in the main paper these alternative options and attempt to add a realistic preservation uncertainty to their data in order to provide a time series with appropriate uncertainty.

Because my issues with the paper are quite fundamental and need addressing before it proceeds further I add only a few less important technical issues below. I should add that the actual analytical achievement is impressive and I have no complaints or concerns about the analytical values or ages used in this paper.

Thank you for the thoughtful assessment of our manuscript. We hope we have sufficiently addressed the important concerns brought up in this review.

Though our comparisons of CO₂ and CH₄ concentrations downcore compared to EDC have led us to infer that the deeper samples represent long-term averages that have smoothed out orbital-scale variations, we recognize that there might be alternative interpretations. In the supplement, we now discuss a range of scenarios associated with the 1) extent of a climate-dependent preservation bias and 2) the extent of physical mixing and/or layer thinning. We present a most likely scenario, but include the range of possibilities for CO₂ histories in Figure 1

(as annotations). By incorporating these other scenarios, we are bounding our record with uncertainty that we believe will satisfy the concerns discussed here. It is important to note that CH₄ concentrations remain quite stable across the Pleistocene, no matter which scenario is used. The new section of the supplement has been pasted below, and can be found starting on line 89 of the new supplementary document.

"Preservation of atmospheric signals in ALHIC1901

ALHIC1901 measurements of CO₂, CH₄, and $\delta^{18}\text{O}_{\text{atm}}$ for samples with $^{40}\text{Ar}_{\text{atm}}$ ages between 400-800 ka recover $42 \pm 1\%$, $38 \pm 1\%$ and $23 \pm 1\%$ of the EDC range for that time period, respectively. These fractions were derived by first interpolating each EDC record¹⁸⁻²⁰ to a 50 year even spacing. We then created 10,000 Monte Carlo simulations of each EDC record, incorporating published analytical uncertainty¹⁸⁻²⁰. We report the fraction that the range in values for each parameter in ALHIC1901 represents relative to the average range of the Monte Carlo simulations. These ranges were calculated using all ALHIC1901 measurements from 400-800 ka, regardless of depth.

When considering how to extrapolate the extent of preservation bias and signal attenuation (due to physical mixing and/or thinning) to measurements older than EDC, we separated the two processes and assessed their individual impacts on the gas records:

1. Preservation Bias

For samples older than the continuous ice core record we examine two endmembers: 1) uniform preservation bias through time, or 2) greater preservation bias in the late Pleistocene than during the late Pliocene/early Pleistocene. The latter scenario would arise if the glacial-interglacial accumulation contrast was greater in the late Pleistocene, a reasonable assumption considering that early Pleistocene and late Pliocene glacial minima were likely less extreme. Under this assumption, values measured at 2.9-2.7 Ma would represent unbiased long-term atmospheric averages. Comparing the mean concentrations for CO₂ and CH₄ (250 ppm and 507 ppb, respectively) to the atmospheric averages during the late Pleistocene (220 ppm and 503 ppb, respectively) indicates a CO₂ decline of 30 ppm and no significant change in CH₄.

Evidence for a change in preservation bias may be observed in the $\delta^{38/36}\text{Ar}$ ratio, which records changes in the height of the diffusive zone within the firn column via gravitational enrichment. Most if not all polar ice cores show a glacial-interglacial difference in firn thickness. The value of $\delta^{38/36}\text{Ar}$ is relatively constant throughout our whole record, indicating little change in the gradient between glacial and interglacial layer thickness (Supplementary Fig. 4). The stability in $\delta^{38/36}\text{Ar}$ would be consistent with a preservation bias that is constant throughout the record, though it does not provide definitive proof thereof.

Lastly, though a preservation bias toward glacial periods is not observed in the younger part of our record, we acknowledge that it is possible due to differing ice rheology leading to a change in ice flow between interglacial and glacial layers (Durand et al. 2007). Evidence against this hypothesis comes from the expected cooling trend across the Pleistocene observed in both the water isotope record (Shackleton et al. in review) and mean ocean temperature record (Shackleton et al. in review) from the same ice core. In addition, we would expect to see a difference in $\delta^{38/36}\text{Ar}$ if there was a switch in preservation bias from glacial layers in the early Pleistocene toward interglacial layers in the late Pleistocene.

2. Thinning/physical mixing

To examine the impact of assumptions about thinning and/or physical mixing, we consider two endmembers: 1) either the extent of physical mixing and/or thinning is depth dependent, or 2) the whole core has experienced the same amount of signal smoothing. First, we examine whether the extent of smoothing changes with depth. Plotting the gradient in age vs depth ("age density") against depth shows a change around 140 m, below which the amount of time represented by one centimeter of ice increases by an order of magnitude (Fig. 3 inset). A change in variance of CO_2 vs CH_4 also occurs at approximately this depth (141 m) (Fig. 2d). Ice younger than 800 ka but stratigraphically below this depth threshold records much less variability in all climate properties. The transition coincides with both the appearance of ice older than 1 Ma and evidence for contamination of trapped CO_2 by respired organic carbon (~141 m depth). We can replicate this change in signal attenuation by employing different smoothing above and below 141 m (Gaussian windows differing by an order of magnitude; Fig. 2b-c). Because we believe the preservation bias is uniform downcore, we again used the synthetic EDC record weighted to S27 layer thickness as described in the main text. We generally are able to reproduce the observed distributions above and below 141 m ("short" vs "long" snapshots). $\delta^{18}\text{O}_{\text{atm}}$ does not show an interglacial bias in the same way as CO_2 and CH_4 which is similar to what we see in the Allan Hills record (Fig. 2). This is likely because the 800 ka $\delta^{18}\text{O}_{\text{atm}}$ record resembles precessional variations²⁰⁻²² whereas CO_2 and CH_4 are dominated by 100 kyr periodicity in the late Pleistocene.

Results show a slight disagreement between the model distribution of CO_2 and the ALHIC1901 measurements above 141 m (Fig. 2b). However, considering the much shorter amount of time likely represented by the shallower samples, a much larger dataset than we have available would likely be required to be certain of this result. Furthermore, if the stratigraphy above 141 m is dipping, we would expect to measure less variability across depth. Thus, it is possible that a higher resolution dataset in this shallow section would exhibit a more representative distribution of EDC. Alternatively, the larger variance in CO_2 relative to CH_4 and $\delta^{18}\text{O}_{\text{atm}}$ could partially be explained by small deviations from true atmospheric CO_2 values. A number of processes have been observed to alter the CO_2 concentration in ice cores, such as the presence of melt layers, carbonate or silicate weathering, oxidation of organic carbon, and post-coring effects²³⁻²⁷. Apart from silicate weathering and carbonate precipitation, most known processes increase the CO_2 concentration in bubbles leading to an overestimate of atmospheric CO_2 . While in some cases CO_2 can be highly enriched (for example, due to oxidation of organic matter near the bedrock interface), in other cases alteration is subtle, with differences between ice cores reported in the 5-20 ppm range^{8,19,28}.

The change in age density in ALHIC1901 indicates that samples below 141 m (which include all samples older than 1 Ma) may have experienced roughly 10 times more signal attenuation than the samples above 141 m. With each sample below 141 m representing more time, a smaller sample size may achieve a more representative distribution, which may explain why we see good agreement between the experimental distributions and the deep Allan Hill samples (Supplementary Fig. 2c). Thus, there appear to be at least two types of paleoclimate snapshots in our record: 1) short-exposure snapshots in shallow ice that preserve some of the true glacial-interglacial variability, and 2) long-exposure snapshots in near-basal ice that preserve climate signals that more closely represent the mean over a glacial cycle. Examining just the long-exposure snapshot record reveals a 20 ppm decline in CO_2 from 2.9-1.2 Ma followed by stable CO_2 concentrations through the MPT and into the late Pleistocene, and no change in CH_4 (Fig. 1). Considering the sample number gets smaller with increasing age in the current dataset (late Pliocene samples: $n = 2$), we also analyzed the long-exposure snapshot record by binning samples in time and conducting statistical t-tests assuming unequal variance. The mean CO_2 and CH_4 values were significantly different (p -value < 0.05) between the pre-MPT (>1.5 Ma, $n=10$) and post-MPT (<0.8 Ma, $n = 16$) samples (Supplementary Fig. 5), suggesting a drop in CO_2 and CH_4 of 10 ppm and 10 ppb between these two binned time periods, respectively.

As an alternative approach, we can assume the time represented by each sample is the same for the entire record. Conducting a t-test between early Pleistocene (older than 1.5 Ma) and late Pleistocene (400-800 ka) samples shows lower CO₂ in the older samples and an insignificant difference in mean CH₄ (Supplementary Fig. 5). Under this endmember scenario, CH₄ shows no change from 2.9-0.8 Ma while CO₂ exhibits a ~7 ppm increase.”

Figure 1 in the main text has been edited to include these scenarios as annotations, producing uncertainty on the secular decline observed across the Pleistocene. The main text has been edited to include this uncertainty as well, stating the long-term change in CO₂ observed is from +7 to -30 ppm from 2.9 to 0.8 Ma. While we wholeheartedly agree that incorporating uncertainty into our record is important, we want to emphasize that it is extremely unlikely, based on our current knowledge of gas preservation in ice, that CO₂ was as high in the early Pleistocene as indicated by average estimates from boron isotope proxy reconstructions.

Minor technical issues

Line 83. Shackleton submitted. I am assuming this is not the parallel MOT paper but rather a related paper about Miocene and Pliocene ice. I understood that Nature doesn't allow unpublished work to be referenced so this will have to depend whether that other paper is published before this one.

We recognize that without being able to see one of the papers originally submitted as a third companion paper, some fundamental details on the setting were lost. The paper referred to here (now Shackleton et al. 2025) is currently in review at PNAS and publicly archived on ESS Open Archive (DOI: 10.22541/essoar.173897232.24923969/v1)

Short and long snapshots. It would have been very helpful if the data table (xls file) had a column saying which samples are used as long and which as short snapshots. This could also have been more clearly defined in the text.

Thank you. The data table has been updated to include a column for “short” and “long” snapshots. We also hope that the more detailed discussion of the extent of physical mixing/thinning as well as the addition of an age density plot provides more support for the short vs. long snapshot argument (new main text Fig. 3)

Figure 1. I believe this is only the data that have NOT been corrected using ¹³C. Please clarify this in the caption. I have to say that the correction, while clever and interesting, is a bit of a distraction, since none of the conclusions of this paper rely on any of the

corrected data. I would recommend that the authors use the main text to explain their ideas about averaging properly and take most of the discussion of correcting CO₂-rich data into the supplement or a separate paper.

The caption for Fig.1 has been updated to clarify that only pristine samples were plotted. Corrected CO₂ measurements are only plotted in Fig. 4 of the main text and the plot of $\delta^{13}\text{C}$ vs $1/\text{CO}_2$ has now been moved to the supplement (new supplementary Figure 8, former main text figure 2). We have now placed more emphasis on the various interpretations of pristine samples and their impact on the CO₂ record. That said, we believe the results of our corrected CO₂ are important to discuss as they extend the record to 3.1 Ma and generally agree with the conclusions from our pristine record.

Line 156 “our CO₂ concentrations across the MPT agree with the CO₂-leafwax record”. This is way too strong a statement based on Fig S7A. Sure the concentrations are similar but the the long term smoothing of the leafwax data contains a pattern I can’t see in the ice core data.’

This line has now been changed to “our CO₂ concentrations across the MPT fall within the range observed in the CO₂-leafwax record (Supplementary Fig. 9).” The line is now 190-192 in the main text.

Supplement line 91-93, do these percentage variances refer to only the short snapshots? Please clarify.

These percentages refer to all 400-800 ka aged ALHIC1901 samples. This line has been added to the supplement (line 95): “These ranges were calculated using all ALHIC1901 measurements from 400-800 ka, regardless of depth.”

The organization of our new draft starts with a discussion of all measurements that overlap with EDC, followed by a discussion of preservation bias and signal attenuation from physical mixing/thinning. We use the record of $\delta^{38}\text{Ar}/^{36}\text{Ar}$ to determine that whatever preservation bias exists in the shallower section is most likely uniform downcore (New supplementary Figure 4), and the change in the record is dominated by a change in signal attenuation, which seems to happen abruptly at 141 m as evidenced by the order-of-magnitude change in age density as well as the change in variance in CO₂ and CH₄.

Response to Referees

Referee #1 (Remarks to the Author):

Many thanks for this second version of the manuscript. I now understand that the dating discussion is in another manuscript which is available, it helps a lot.

These unique data are very important and should be published. I still have some questions on the method and I found that there was a lack of information some time. I detail the comments below.

Thank you for the feedback on our manuscript. Below we have provided point-by-point responses to your comments.

1- Summary paragraph :

Why do you speak about « secular changes » here ? Why do you use the adjective « secular » ?

We use the term “secular” to discuss the trend seen in the ‘long’ snapshots because we interpret these samples to represent averages over one or more glacial cycles. This word has been replaced with “long-term” to reduce confusion.

2- L. 81 « individual folds and age discontinuities are resolved » -> Are we 100% sure of it ? How can we exclude that some folds or hiatus can still exist at smaller scale ?

It is possible that age discontinuities exist at a smaller scale than we can measure using the ^{40}Ar geochronometer. While we are limited by the required sample size for this method, we did subdivide 6 samples along the length of a single ^{40}Ar sample and measured methane and carbon dioxide concentrations at 2-3 cm resolution (New Figure S1). We do not observe large variance in these samples, providing evidence that a single ^{40}Ar age is not a mixture of different ages. These results were not included in the original submission and are the product of work done by Demetria Eves who has now been listed as a co-author. These measurements are now included in the data table submitted along with our manuscript.

We now mention these subsamples in the main text:

NEW LINE 72: Third, by increasing the density of age determinations (79 samples in the bottom 25 meters), individual folds and discontinuities are resolved¹³. Of the 79 samples analyzed, 54 are older than 800 ka, with ages ranging from 800 ± 88 ka to 4000 ± 440 ka. While we are limited by the same size requirement of the $^{40}\text{Ar}_{\text{atm}}$ age technique, we do not observe large variance in CO_2 or CH_4 when measured in 2-3 cm resolution along the length of a single $^{40}\text{Ar}_{\text{atm}}$ sample, providing evidence that a single ^{40}Ar age is not a mixture of different ages (Figure S1).

Also, although we are limited by the sample size required for determining an age with the ^{40}Ar method, we emphasize that the high-resolution age scale produced for the ALHIC1901 ice core is an improvement from previous studies.

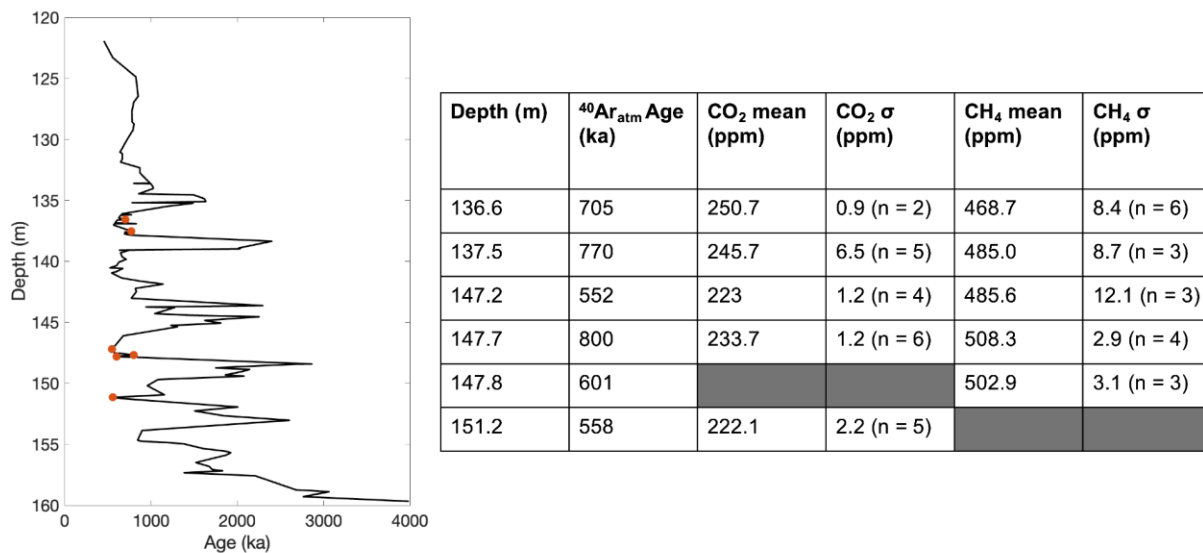


Figure S1. Left: Location of six samples (plotted as orange circles on the ALHIC1901 chronology) subdivided along the length of a single $^{40}\text{Ar}_{\text{atm}}$ sample and analyzed for CO_2 and CH_4 concentrations at 2-3 cm resolution. Right: Table showing the mean and standard deviation of each subset of CO_2 and CH_4 samples. The low variability exhibited in these samples implies that single $^{40}\text{Ar}_{\text{atm}}$ measurements are likely not mixtures of ice of different ages.

3- L. 103 – I do not really understand the concept of « age density » . We can not exclude hiatus in the record, especially below 140 m so how would age density be defined in this case ? Also it is never mentionned that sections of ice are missing in this ice core. It is always about thinning and folding. Some thoughts about hiatus should be introduced.

We introduced the concept of “age density” in order to visualize and approximate the extent of signal attenuation (due to some combination of thinning and folding) throughout the ice core. This is calculated as the change in time over the change in depth (dt/dz). In this comment, I believe the reviewer is questioning the interpretation of age density in the bottom 20 m if there are missing parts of the record. For example, if the chronology shows a jump from 800ka to 2 Ma, we would calculate a large age density. However, if there are sections of ice from a large time interval, missing, say from 1-1.5 Ma, then age density would appear larger than it actually is. As an extreme endmember scenario, if the change in age density was entirely caused by hiatuses then there would be no distinction between ‘long’ and ‘short’ snapshots and all measurements would have indistinguishable time resolution. In our manuscript we discuss this as an alternative interpretation of our record (treating all measurements as indistinguishable with no short vs long categorization). In the latest version of the manuscript, we now mention the possibility of hiatuses in conjunction with the discussion of this alternative interpretation. The new text describing

this issue can be found under the section “Alternative interpretations” in the supplement and pasted below:

NEW LINE S163: *We also must consider the possibility that the time resolution of each sample is the same, regardless of depth. This would be the case if the drastic increase in age density that we observe at 140 m is an artifact of large sections of the record being missing rather than a change in time resolution in a continuous but compressed record. While this seems an unlikely scenario, to examine the implications we conducted a t-test between early Pleistocene (older than 1.5 Ma) and all late Pleistocene (400-800 ka) samples (‘short’ and ‘long’ snapshots). Results show lower mean CO₂ in the older samples and an insignificant difference in mean CH₄ (Figure S12). Under this endmember interpretation, mean CH₄ shows no change from 2.9-0.8 Ma while CO₂ exhibits a small, ~7 ppm increase.*

4- Related to the previous comment, it is important to understand what you mean by « climate dependent preservation bias ». Is it that you have more hiatus during certain period ? Or more diffusion / thinning during a certain climate ?

Thank you for this question - we have now included a definition for climate-dependent preservation bias in the main text:

NEW LINE 63: *We have applied methods to evaluate the effects of signal attenuation and potential sampling bias in warmer vs. colder climate periods (“climate-dependent preservation bias”) in each snapshot.*

5- The term « physical mixing » is not clear. What do you mean ? A mixing of two layers of ice because of very small scale folding ? Or diffusion of the tracers within the ice among adjacent layers ? Please further explain what physical process you have in mind.

We have adjusted our wording to signify that signal attenuation in the core could be from any type of post-depositional process such as thinning, folding, and recrystallization at various length-scales during flow from the source region (**NEW LINE 97**).

Regarding diffusion, we believe molecular diffusion is not a dominant process impacting our record because 1) abrupt changes are observed in the water isotope and age records from co-registered samples, and 2) the model constructed in Yan et al. 2019 shows diffusion is likely to have a minimal effect. This is discussed in the main text as follows:

NEW LINE 136: *Additional smoothing by molecular diffusion is one possibility, though previous modeling work shows that this is unlikely for these timescales at the cold temperatures of the Allan Hills¹⁵.*

6- From l. 126 (and despite the supplement), it lacks explanations and more description with additional figures (can be in the supplement).

a. I can not understand how you create a synthetic 400-800 ka ice core record using glacial – interglacial thickness variations from the text and figure S2 which is not very useful as it is (we do not see the glacial – interglacial cycle). I suggest to simply show the synthetic curve and explicitly give the formula used to produce it.

Thank you for the feedback. We now include a more detailed section in the supplement that describes how the synthetic record was developed (pasted below).

NEW LINE S99: *To explore the effects of a climate-dependent preservation bias, we created a synthetic ice core record that has glacial-to-interglacial annual layer thickness variations we would expect at Allan Hills. We limited the synthetic record to 400-800 ka to overlap with ALHIC1901 measurements within $^{40}\text{Ar}_{\text{atm}}$ age uncertainty. We first calculated the annual layer thicknesses of EDC and the S27 ice core from Allan Hills using their respective chronologies^{22,23} (after evenly spacing the records with respect to time). Comparing the two records of annual layer thickness revealed an order-of-magnitude difference (Figure S8). We thus exaggerated the EDC layer thickness pattern to produce a synthetic depth record for EDC with a more extreme difference between glacial and interglacial layers. This was done by first identifying the peaks and troughs of the original EDC layer thickness record and interpolating between them using the nearest neighbor method to produce a simplified layer thickness record with a square wave pattern (Figure S7). We varied the amplitude of the simplified square-wave record, increasing the relative interglacial to glacial layer thickness contrast by a factor of 10 to match the layer thickness record of the S27 ice core. This was done by multiplying the peak and trough values of the simplified record by 5 and 0.5 respectively (Figures S8). We then interpolated the EDC gas records^{18–21} onto this synthetic chronology, producing hypothetical continuous records for the Allan Hills BIA. We also include the results of a layer thickness contrast 3 orders of magnitude larger than the simplified square-wave record (multiplying the peak and trough values by 50 and 0.05, respectively), providing an extreme preservation bias scenario. The difference in distributions for each gas parameter between the simplified EDC layer thickness record and the exaggerated records can be seen in Figure S5.*

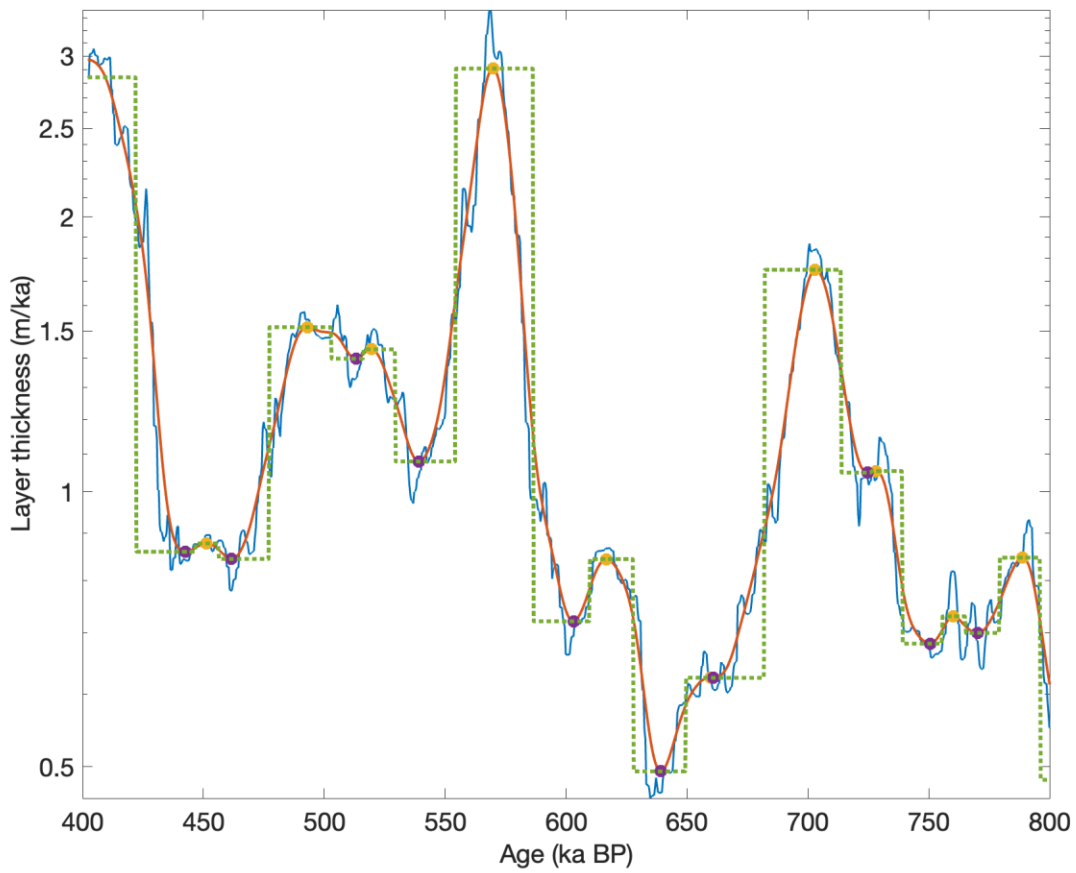


Figure S7. EDC layer thickness (blue curve) calculated from the AICC2023 chronology²² and the smoothed EDC layer thickness curve (red) using a Gaussian-weighted moving average with standard deviation of 4000 years. Circles show peak (yellow) and trough (purple) values identified in the smoothed EDC record, and the green square wave is the simplified layer thickness curve constructed from interpolating between peak and trough values using the nearest neighbor method.

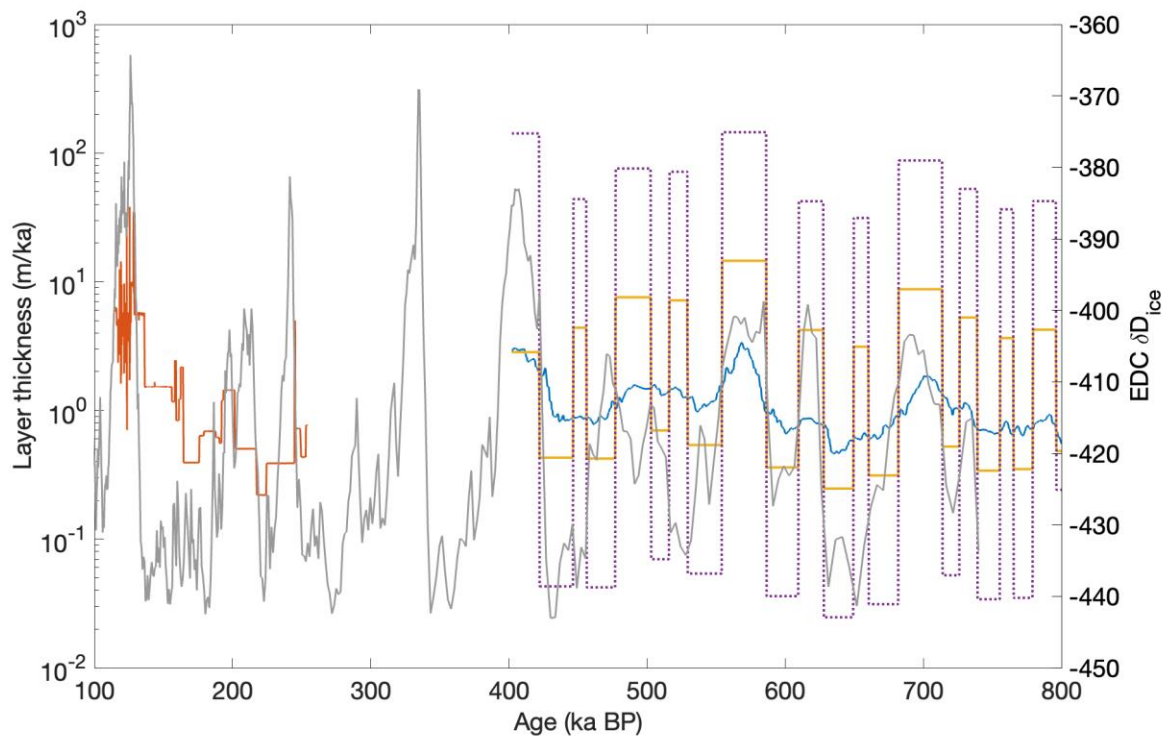


Figure S8. EDC layer thickness (blue curve) and S27 layer thickness (orange) calculated from their respective chronologies^{22,23}. The yellow and purple curves are synthetic, simplified layer thickness records where the difference between peaks and troughs from the original EDC layer thickness are exaggerated by one and three orders of magnitude larger, respectively. Layer thicknesses are plotted on a log scale. The water isotope record from EDC⁴⁸ is overlaid. Peaks and troughs of the layer thickness record generally align with the glacial cycles observed in the water isotope record.

b. « Thinning and small-scale folding were simulated by applying a gaussian smoothing. » So if I understand well, you do not apply ice diffusion ? Still, ice diffusion is very important for ice older than 1 million year and should be considered (Bereiter et al., 2014). This effect should be implemented.

Applying diffusion to our model would require we make assumptions about the flow path and consequently the temperature history of the ice. From our current understanding, we believe molecular diffusion is not a dominant process impacting our record because 1) abrupt changes are observed in the water isotope and age records from co-registered samples, and 2) the model constructed in Yan et al. 2019 shows diffusion is likely to have a minimal effect. Please refer to our response to comment #5 to see where in the manuscript we discuss this topic.

c. « Applying a gaussian smoothing » -> with what parameters ? Please give the detail of this gaussian smoothing (standard deviation of the distribution). It is important that the details of the calculations are given to follow what is done.

This information was provided in the Figure 2 caption: **(2a). Magenta distributions in 2b and 2c come from smoothed versions of the weighted EDC depth record and differ by an order of magnitude (Gaussian-weighted moving average with standard deviation of 20 m (2b) and 200 m (2c)).**

However, in this new manuscript we have decided to exclude this figure because we think it was an overly complicated way of discussing signal attenuation in the core. We think this is most clearly represented in the cross plot in Figure 1d and the age density plot (Figure S4).

d. I am surprised by the sentence beginning l. 145. The diffusion for O₂, CO₂ are given in the paper by Bereiter et al. (2014) so it is possible to infer the effect of diffusion for both CO₂ and d18O_{atm}.

This is the text being referred to in this question: **Additional smoothing by molecular diffusion is one possibility, though Yan et al. 2019 argue that this is unlikely for these timescales at the cold temperatures of the Allan Hills. If important, this process could be contributing to the narrower $\delta^{18}\text{O}_{\text{atm}}$ distribution as we expect higher diffusivities and more smoothing associated with O₂ due to its smaller molecular diameter compared to CO₂ and CH₄^{35,36}.**

Yan et al. 2019 did extensive modeling of molecular diffusion for nearby Allan Hills ice cores, which we reference in our discussion of the impact this process may have on our records. At this time, the model from Yan et al. 2019 is our best understanding of molecular diffusion on this ice. While we believe this process to have minimal impact on our ice and gas records from ALHIC1901, we referenced Bereiter et al. 2014 to support our argument that molecular diffusion would impact d18O_{atm} more than CO₂.

7- I am not fully convinced by the d38/36Ar argument against gas loss (l. 156). As d18O_{ice} is varying with time, d38/36Ar could also vary with time because of changes in accumulation rate and temperature. In parallel, you mention O₂/N₂ measurements in the supplement which you could use for this purpose (and you actually used for correction of the gas loss effect – so there is some gas loss effect ? This is not clear). Why not showing and using them ?

This is the text being referred to here: **Measurements of $\delta^{38/36}\text{Ar}$ ratios of trapped air, an indicator of firn thickness from mass-dependent fractionation due to gravitational settling and/or gas loss, do not vary systematically with sample age, suggesting that whatever interglacial bias we observe in the discontinuous young (<800 ka) ice does not change with sample age.**

In this section, we are not using d38/36Ar as an indicator of gas loss and probably confused the reader by including it in this sentence. Gas loss in our measurements is corrected for using dO₂/N₂ and is not relevant to the discussion of how d38/36Ar changes with depth or time as it is discussed here.

a. there is a significant difference between the yellow and magenta distribution for CO₂ (2b) – which is describe as « slightly different » in the text but it is not really « slight » - at least give the exact number of the difference in the text.

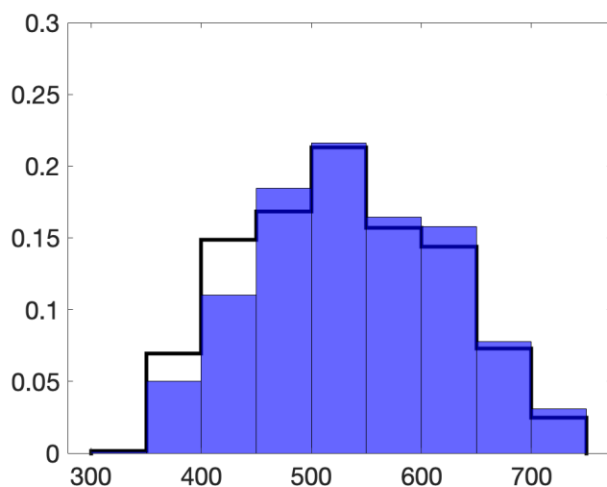
We have adjusted the wording in the main text to distinguish between the good agreement for all other parameters and the short snapshot CO₂ record:

In the main text, the whole section titled **Interpreting Allan Hills ice** is relevant to this question. We have specifically pasted a new paragraph here:

NEW LINE 145: *Finally, the much larger discrepancy between ‘short’ and ‘long’ snapshot means in CO₂ compared to CH₄ or $\delta^{18}\text{O}_{\text{atm}}$ could suggest a small amount of excess CO₂. A number of processes have been observed to alter the CO₂ concentration in ice cores, such as the presence of melt layers, carbonate or silicate weathering, oxidation of organic carbon, and post-coring effects^{23–27}. Apart from silicate weathering and carbonate precipitation, most known processes increase the CO₂ concentration in bubbles leading to an overestimate of atmospheric CO₂. While in some cases CO₂ can be highly enriched (for example, due to oxidation of organic matter near the bedrock interface), in other cases alteration is subtle, with differences between ice cores reported in the 5-20 ppm range^{28–30}. If present, the amount of enriched CO₂ is too small (5-20 ppm) to correct for considering existing uncertainties in the isotopic endmembers of the CO₂ mixing model (see Methods).*

b. I have difficulty to understand why there is no magenta rectangles below 450 ppbv on figure 2b (CH₄) while the difference between the blue and black rectangle on ffigure 2a are not so different.

Thank you for the clarification question on Figure 2. The blue distribution in Figure 2a for CH₄ was incorrectly plotted. The correct one (pasted below) exhibits a slightly bigger difference between the EDC distribution (black) and the synthetic distribution (blue). We do not see glacial-level values in any of the three gas parameters as a result of applying a gaussian smoothing to the EDC record.



That said, we think our experiment may have been unnecessarily complicated and have simplified the results and replaced the panels in Figure 2 with a new figure (Figure S5). This compares distributions of each gas parameter after exaggerating the EDC layer thickness by varying orders of magnitude. We also plot the mean value of each distribution as squares and compare these to the means of the short and long snapshots of ALHIC1901. We still include the experiment where we randomly sampled from the synthetic distributions after applying a gaussian smoothing with varying kernel sizes. These results are now in the supplement (NEW Figure S5)

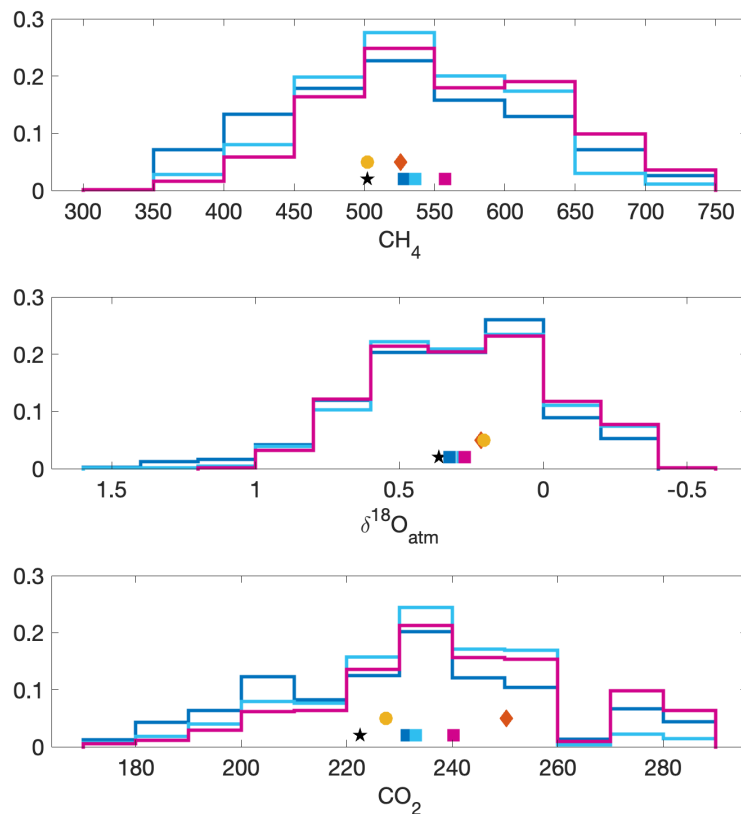


Figure S5. Normalized histograms of CH_4 , $\delta^{18}\text{O}_{\text{atm}}$, and CO_2 (from top to bottom) of EDC from 400-800 ka (dark blue). Additional histograms show the change in distribution of each gas property after exaggerating the amplitudes of the EDC layer thickness by one (light blue) and three (pink) orders of magnitude (see Methods). Squares plotted below the histograms represent the mean value of each gas property for the various distributions. The time-averaged mean of EDC from 400-800 ka is plotted as a black star. Mean values of ALHIC1901 measurements are plotted as diamonds ('short' snapshots) and circles ('long' snapshots).

9- Supplementary material

a. l. 49 : show the data or give some order of magnitude of the corrections which are done.

We are not sure what this comment refers to, but believe it is the $\delta^{13}\text{C}$ -corrected CO_2 data? Both the raw and corrected data are included in the data table submitted along with the manuscript.

b. L. 141 : what do you mean by a preservation bias which is « uniform downcore ». It would be helpful to have the equations to understand/follow what you mean for all the section on the preservation bias.

We have rewritten the supplement to clarify what we mean by preservation bias and it being uniform downcore:

New Line S121: *Evidence for a change in preservation bias might be observed in the $\delta^{38/36}\text{Ar}$ ratio, which records changes in the height of the diffusive zone within the firn column via gravitational enrichment. Most if not all polar ice cores show a glacial-interglacial difference in firn thickness. The value of $\delta^{38/36}\text{Ar}$ is relatively constant throughout our whole record, indicating little change in the gradient between glacial and interglacial layer thickness (Figure S9). The stability in $\delta^{38/36}\text{Ar}$ would be consistent with, but does not prove, a preservation bias that is constant throughout the record.*

This can also be found in the main text:

NEW LINE 125: *An additional explanation for the difference between ‘short’ and ‘long’ snapshots could be a change in climate-dependent preservation bias that could cause a shift in the mean value for each gas property. However, results from a statistical resampling of each EDC ice core record after weighting them to exhibit more extreme differences between glacial and interglacial layers (see Methods, Figures S7 and S8) shows that preservation bias has a small effect on the mean value for each gas property (<20 ppm and <50 ppb for CO_2 and CH_4 , respectively, Figure S5). Furthermore, measurements of $\delta^{38/36}\text{Ar}$, an indicator of firn thickness and accumulation rate, does not vary in our samples (Figure S9), suggesting that the climate-dependent preservation bias did not change and therefore is not likely to explain the different mean values between short and long snapshots.*

c. You give different explanation of the $\delta^{18}\text{O}_{\text{atm}}$ not having an interglacial bias in the main text (diffusion effect) and in the supplement (precessional signal). Check coherency

Thank you for bringing this to our attention - we believe the narrow distribution in $\delta^{18}\text{O}_{\text{atm}}$ and the lack of an interglacial bias could be explained by either or both of these processes and have made sure both are written where $\delta^{18}\text{O}_{\text{atm}}$ is discussed:

In the main text:

NEW LINE 136: *The very narrow distribution observed in $\delta^{18}\text{O}_{\text{atm}}$ is not fully explained by our analysis. Additional smoothing by molecular diffusion is one possibility, though previous modeling work shows that this is unlikely for these timescales at the cold temperatures of the Allan Hills¹⁵. If important, this process could be contributing to the narrower $\delta^{18}\text{O}_{\text{atm}}$ distribution as we expect higher diffusivities and more smoothing*

associated with O_2 due to its smaller molecular diameter compared to CO_2 and CH_4 ^{18,19}. Alternatively, the narrow distribution may be because the 800 ka record of $\delta^{18}O_{atm}$ follows precession^{20–22} whereas CO_2 and CH_4 have large variability on 100 kyr time scales in the late Pleistocene.

Referee #2 (Remarks to the Author):

I thank the authors for their revised paper. Several issues have been explained more clearly and some interpretations have been toned down. The paper, with its improved dating compared to previous papers, and its use of direct ice core greenhouse gas measurements, remains a potential breakthrough that most likely warrants publication in Nature.

There remains one area of the paper where I feel the authors may not have fully understood what my concern was on the last round. This concerns the preservation bias. The authors go to considerable (perhaps too much) length to explain that the deep samples cover a very long period of time. I felt this went without saying, and myself I feel that Figure 3 does not deserve to be in the main text. However, this does not cover the issue about preservation bias, which has nothing to do with how much time each sample covers. The question is whether each sample averages across time in a consistent way. Looking at Fig 1a and b it's obvious, as stated in the paper, that there is a strong interglacial bias for the short snapshots. If one was to average all the samples from the short snapshots, the average would not be a time average of the EDC reference value (in fact Fig 2b, CO_2 panel, with a strong mismatch between the measured yellow and the magenta calculation, suggests that the interglacial bias is even stronger than the authors have used). However many of those short snapshots you take, if you simply average them without a correction, they will give you a value of around 250 ppm, not the EDC time average of around 220 ppm. This will still be true if you combine them by joining them together and thinning them, which is what a long snapshot is.

The authors suggest that the interglacial bias is due to the higher accumulation rates in the interglacial, ie interglacial time is way overrepresented in any length of ice. They also claim (partly from the data in Fig S4) that the interglacial bias persists throughout their record. But this cannot be true: when it comes to the long snapshots for the same time period, they do give roughly the EDC time average, implying that the interglacial ice is not overrepresented. Thinning and folding do not in themselves alter this. So what has happened – has some kind of differential thinning compensated so that there is (by chance) a time average in these long snapshots.

The reason why this is critical is because it begs the question whether the long snapshots below 800 kyr behave like the long snapshots or the short snapshots, and we simply don't know. At one end of options, the early Pleistocene snapshots may have an interglacial bias, they might have no bias, and they might even have a glacial bias. This is really not well-explained in the text, in fact not at all in the main text. One can partly understand this from Fig S5 but even there the issue is not made clear.

The authors really need to have another go at this and clarify it in the text. Do they have an idea why the interglacial bias is not seen in the long snapshots from 800-500 kyr, as it really should still be there unless another process, not mentioned, has taken place. Their preferred option is that all the long snapshots are unbiased recorders (ie without an interglacial bias), but I can't see why this should be. I am Ok with the options they end up with (although they don't cover the possibility of a glacial bias in the early Pleistocene) but it is not well-explained, nor is it clear why they prefer the no-bias scenario.

Thank you for your thoughtful feedback. Though we wrote an extensive section in the supplement that contains a discussion of 1) how deviated our record is from the true time average (“preservation bias”) and 2) how much time may be represented

by each sample (“thinning/physical mixing”), we agree that we could have better incorporated this work into the main text of the manuscript.

We have rewritten the main text to better summarize the results of modelling preservation bias and now spend more time discussing the issue. Along with this new discussion, we have removed figure 2, believing it was an overly complicated way of discussing signal attenuation in the core. We think this is most clearly represented in the cross plot in Figure 1d and the age density plot (Figure S4).

For the discussion of preservation bias, we have added more detail to the supplement explaining how we built our synthetic record and include 2 new figures to help visualize what we did (Figures S7 and S8). We also have a new figure (Figure S5) that shows how the distribution of each gas parameter is impacted by a change in preservation bias. The main conclusion is that we cannot explain the difference between the mean values of short and long snapshots using a change in preservation bias, indicating that the difference is likely due to a sampling bias in the shallower section making the short snapshot dataset an incomplete representation of the full range that exists in the ice core at shallower depths. Evidence for this hypothesis comes from the larger range in $\delta^{18}\text{O}_{\text{atm}}$ measured in shallower depths from the nearby ALHIC1503 ice core (Higgins et al. 2015). Another possible reason for the higher CO_2 values above 141 m is because of a small amount of CO_2 production from organic matter, adding 5-20 ppm to the atmospheric signal.

Additionally, we agree that there should have been more of a discussion of the comparison between our record and the results from our modeling results, particularly the disagreement between the synthetic distributions and the CO_2 short snapshots. This is hopefully clearer in our main text, as follows:

NEW LINE 92: Interpreting Allan Hills ice cores

Paleoclimate archives in the Allan Hills BIA range from stratigraphically intact, younger, ice that preserves the full range of glacial-interglacial variability (core S27¹⁶) to stratigraphically discontinuous ice that overlaps in age with continuous ice cores from the interior (<800 ka), to discontinuous old (>1 Ma) ice with no continuous ice core analog for comparison (Figure 1b). The loss of stratigraphic continuity is presumably due to thinning, folding, and recrystallization of the ice on various length-scales during flow from its source region. Together these processes result in an increase in the amount of time preserved in a given volume of ice (age density or yr/cm; Figure S4) and a complicated age-depth relationship.

The age densities (yr/cm) of paleoclimate records in Allan Hills ice cores, determined from chronologies developed from trace gas alignment (CH_4 and $\delta^{18}\text{O}_{\text{atm}}$ ^{16,17}) and high-resolution $^{40}\text{Ar}_{\text{atm}}$ analyses, span >5 orders of magnitude, from ~1 yr/cm during MIS5 in S27 to ~100,000 yr/cm in the basal portion of ALHIC1901 (between 140 m and bedrock, Figure S4).

In the deepest, highly compressed ALHIC1901 ice, discrete cm-scale samples likely represent averages over glacial cycles ('long' snapshots). In shallower and younger ALHIC1901 ice, age densities of 1-10 kyr/cm may preserve more (though likely not all) of the true glacial-interglacial variability ('short' snapshots).

CO₂ and CH₄, and $\delta^{18}\text{O}_{\text{atm}}$ records from ALHIC1901 further support a shift in the time resolution of samples below 140 m at this site. We observe a reduced range in paleoclimate properties in 'long' (>140 m) snapshots compared to shallower 'short' (<140 m) snapshots of similar age (Figure 1d). 'Short' snapshots younger than 800 ka span some (but not all) of the observed variability for the last 800 ka. Notably, glacial levels are absent, and the mean value of each gas property is higher than that of EDC records of overlapping age (400-800 ka) (Figure S5). For CO₂ and CH₄, this stands in contrast to the good agreement between average EDC values and <800 ka 'long' snapshots (Figure S5). The change in mean value from 'short' to 'long' snapshots could be due to an under sampling of <800 ka ice from ALHIC1901 due to the relatively limited (i.e. ~6 meter) interval where 'short' snapshots are observed in this core. Samples from the nearby ALHIC1503 ice core included a much wider depth range, and the shallower (<800 ka) samples are characterized by a much wider range in $\delta^{18}\text{O}_{\text{atm}}$ values compared to ALHIC1901 (including elevated values (>1.2‰) consistent with glacial maxima¹⁴) (Figure S6).

An additional explanation for the difference between 'short' and 'long' snapshots could be a change in climate-dependent preservation bias that could cause a shift in the mean value for each gas property. However, results from a statistical resampling of each EDC ice core record after weighting them to exhibit more extreme differences between glacial and interglacial layers (see Methods, Figures S7 and S8) shows that preservation bias has a small effect on the mean value for each gas property (<20 ppm and <50 ppb for CO₂ and CH₄, respectively, Figure S5). Furthermore, measurements of $\delta^{38/36}\text{Ar}$, an indicator of firn thickness and accumulation rate, does not vary in our samples (Figure S9), suggesting that the climate-dependent preservation bias did not change and therefore is not likely to explain the different mean values between short and long snapshots.

The very narrow distribution observed in $\delta^{18}\text{O}_{\text{atm}}$ is not fully explained by our analysis. Additional smoothing by molecular diffusion is one possibility, though previous modeling work shows that this is unlikely for these timescales at the cold temperatures of the Allan Hills¹⁵. If important, this process could be contributing to the narrower $\delta^{18}\text{O}_{\text{atm}}$ distribution as we expect higher diffusivities and more smoothing associated with O₂ due to its smaller molecular diameter compared to CO₂ and CH₄^{18,19}. Alternatively, the narrow distribution may be because the 800 ka record of $\delta^{18}\text{O}_{\text{atm}}$ follows precession²⁰⁻²² whereas CO₂ and CH₄ have large variability on 100 kyr time scales in the late Pleistocene.

Finally, the much larger discrepancy between 'short' and 'long' snapshot means in CO₂ compared to CH₄ or $\delta^{18}\text{O}_{\text{atm}}$ could suggest a small amount of excess CO₂. A number of processes have been observed to alter the CO₂ concentration in ice cores, such as the presence of melt layers, carbonate or silicate weathering, oxidation of organic carbon, and post-coring effects²³⁻²⁷. Apart from silicate weathering and carbonate precipitation, most

known processes increase the CO₂ concentration in bubbles leading to an overestimate of atmospheric CO₂. While in some cases CO₂ can be highly enriched (for example, due to oxidation of organic matter near the bedrock interface), in other cases alteration is subtle, with differences between ice cores reported in the 5-20 ppm range^{28–30}. If present, the amount of enriched CO₂ is too small (5-20 ppm) to correct for considering existing uncertainties in the isotopic endmembers of the CO₂ mixing model (see Methods).

In summary, although aspects of the paleoclimate information contained in Allan Hills ice cores remain poorly understood, our results support an interpretation where the oldest ice preserves ‘long’ snapshots of mean atmospheric CO₂, CH₄, and $\delta^{18}\text{O}_{\text{atm}}$ over one or more glacial cycles. Although this interpretation of the Allan Hills ice core records limits our ability to comment on the amplitude of glacial-interglacial variability in these properties, we believe they provide robust records of mean atmospheric composition for CO₂ and CH₄ over the Pleistocene. Climate dependent differences in accumulation rates may introduce a small interglacial bias, shifting the mean higher than the true glacial-cycle average, however our analysis shows this difference is small (within 20 ppm and 50 ppb, respectively). This bias does not appear to change with sample age and therefore likely does not impact the long-term trend observed in the ‘long’ snapshot record. If a climate-dependent preservation bias increased across the Pleistocene with the growth of larger ice sheets and more extreme glacial minima, the long-term trend exhibited in our record may be dampened by ~10 ppm (see Methods). Alternative interpretations of the long-term trends in our record can be found in the supplement.

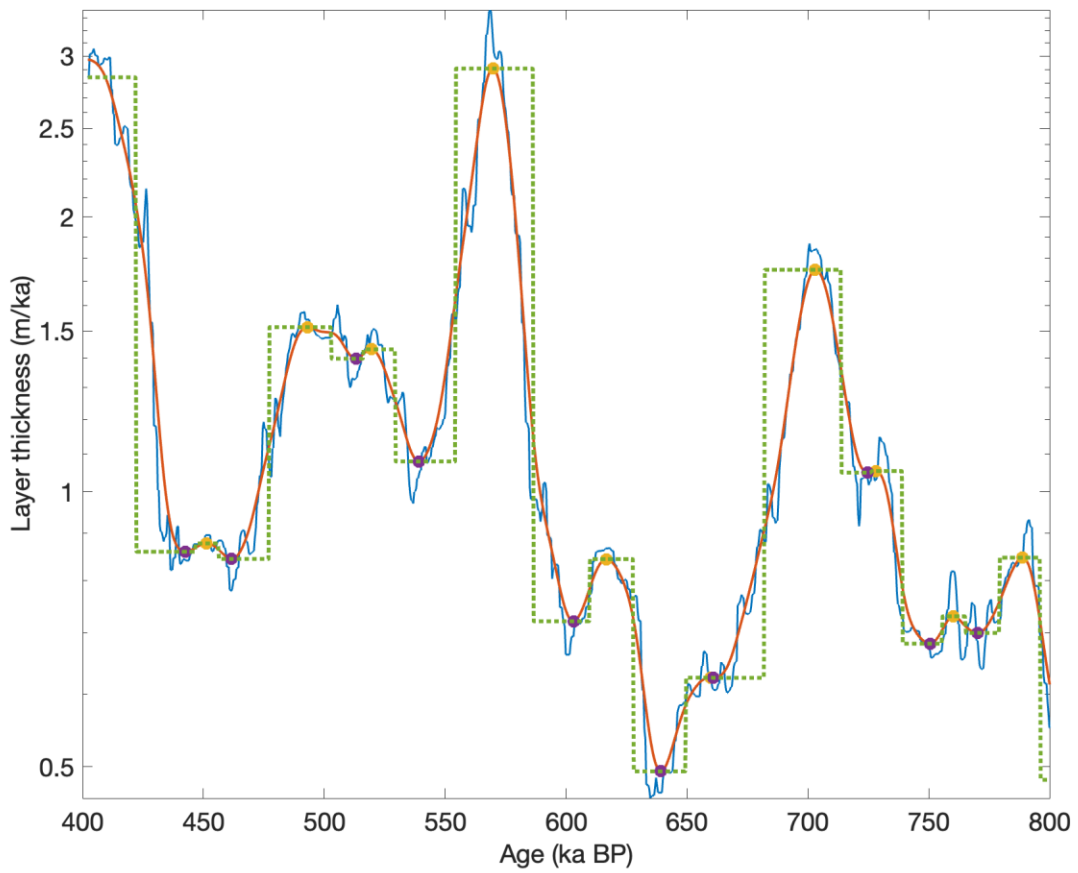


Figure S7. EDC layer thickness (blue curve) calculated from the AICC2023 chronology²² and the smoothed EDC layer thickness curve (red) using a Gaussian-weighted moving average with standard deviation of 4000 years. Circles show peak (yellow) and trough (purple) values identified in the smoothed EDC record, and the green square wave is the simplified layer thickness curve constructed from interpolating between peak and trough values using the nearest neighbor method.

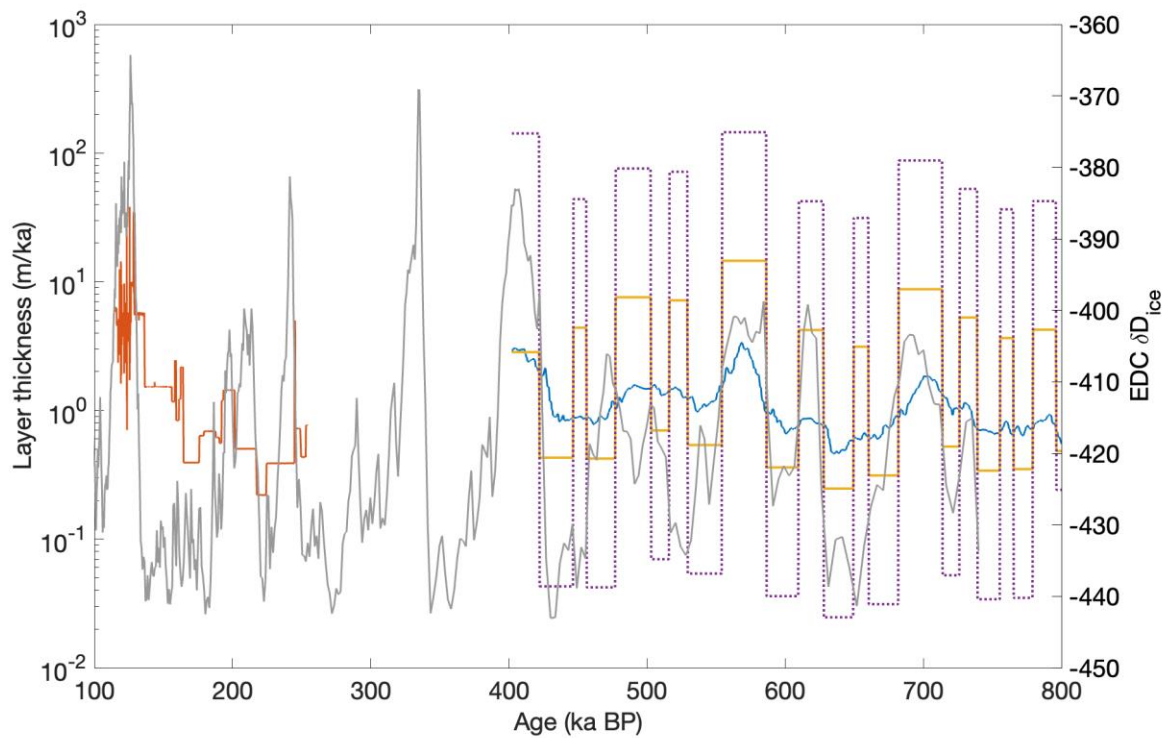


Figure S8. EDC layer thickness (blue curve) and S27 layer thickness (orange) calculated from their respective chronologies^{22,23}. The yellow and purple curves are synthetic, simplified layer thickness records where the difference between peaks and troughs from the original EDC layer thickness are exaggerated by one and three orders of magnitude larger, respectively. Layer thicknesses are plotted on a log scale. The water isotope record from EDC⁴⁸ is overlaid. Peaks and troughs of the layer thickness record generally align with the glacial cycles observed in the water isotope record.

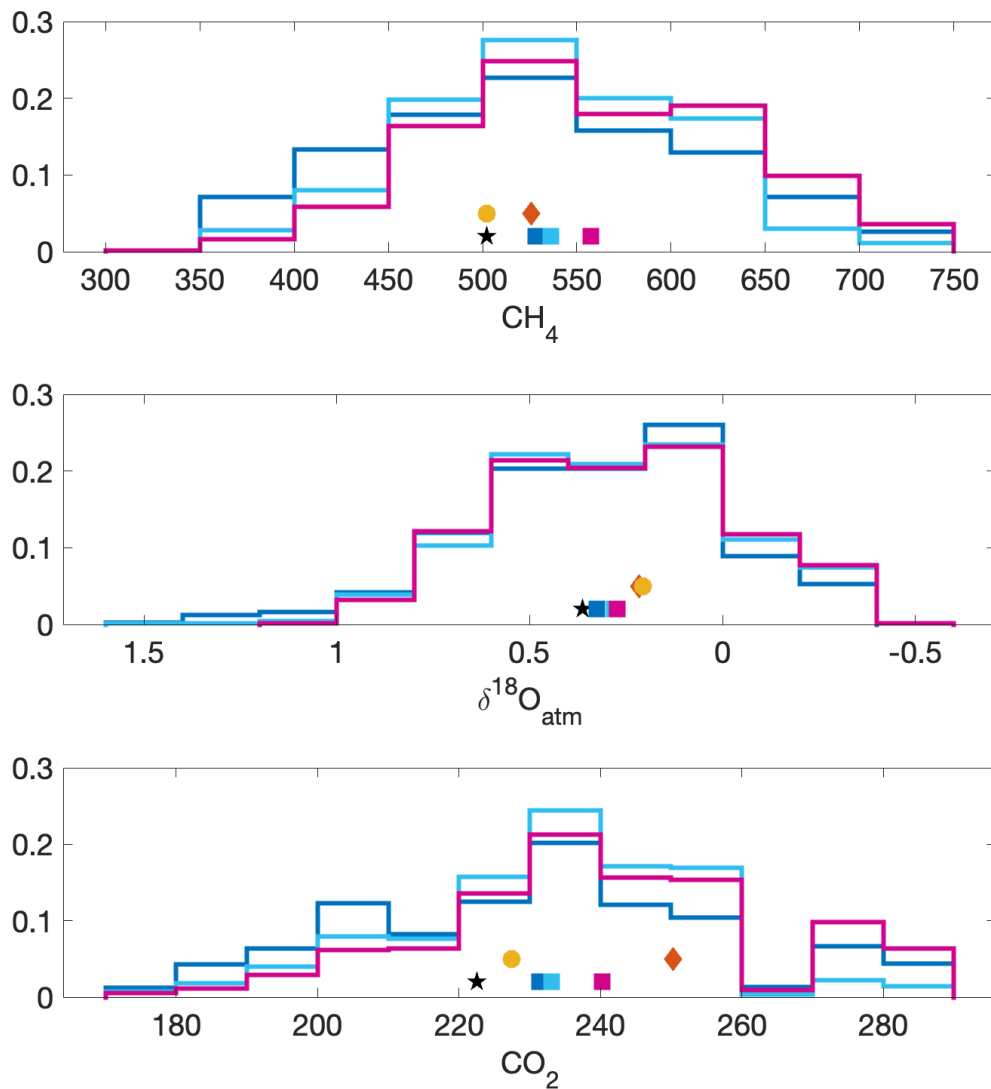


Figure S5. Normalized histograms of CH_4 , $\delta^{18}\text{O}_{\text{atm}}$, and CO_2 (from top to bottom) of EDC from 400-800 ka (dark blue). Additional histograms show the change in distribution of each gas property after exaggerating the amplitudes of the EDC layer thickness by one (light blue) and three (pink) orders of magnitude (see Methods). Squares plotted below the histograms represent the mean value of each gas property for the various distributions. The time-averaged mean of EDC from 400-800 ka is plotted as a black star. Mean values of ALHIC1901 measurements are plotted as diamonds ('short' snapshots) and circles ('long' snapshots).

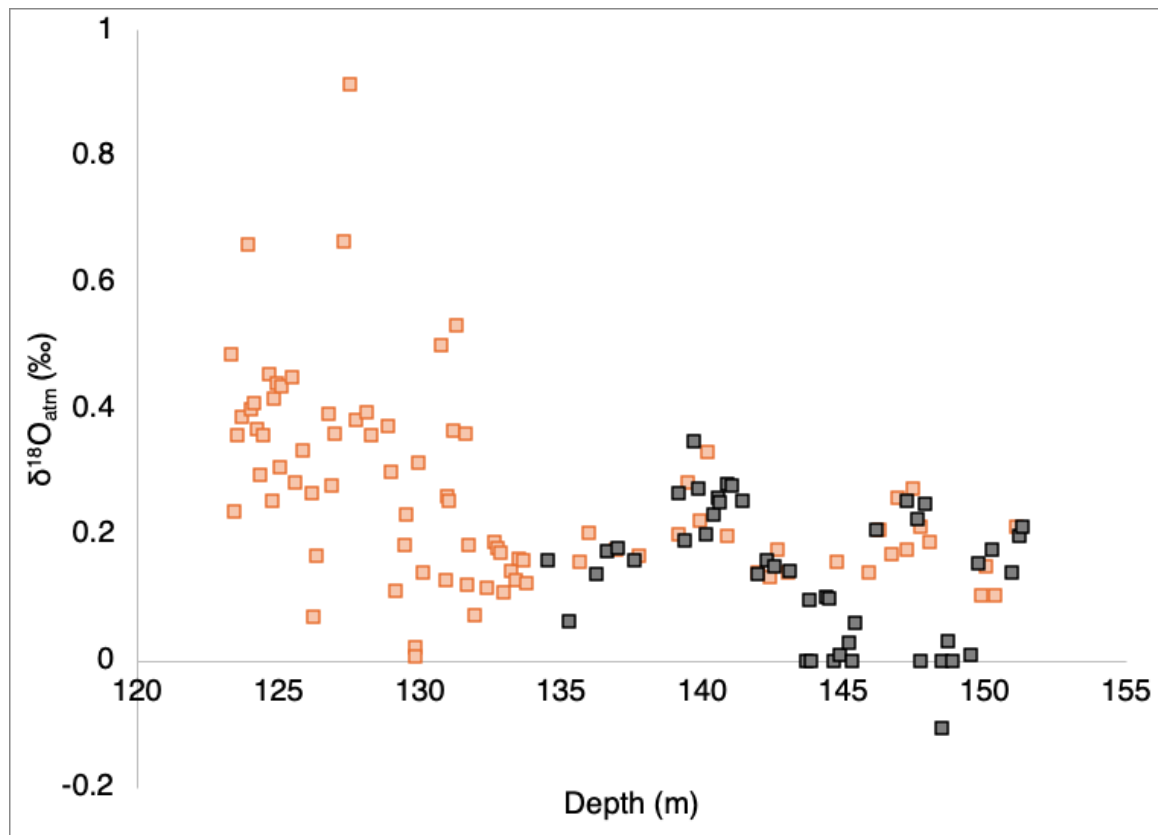


Figure S6. $\delta^{18}\text{O}_{\text{atm}}$ plotted on depth for ALHIC1901 (black squares) and ALHIC1503 (orange squares). ALHIC1503 measurements⁴⁷ are plotted on original depth (have not been adjusted to align with ALHIC1901 measurements). Larger variability in $\delta^{18}\text{O}_{\text{atm}}$ is seen in shallower sections of the ice column than samples measured in ALHIC1901.

A few more detailed comments:

Abstract, l17-19. As I say this seems like a reasonable conclusion but it is really not supported by what is in the text.

[We hope our new description in the main text provides the context for this line in the abstract:](#)

NEW LINE 157: *In summary, although aspects of the paleoclimate information contained in Allan Hills ice cores remain poorly understood, our results support an interpretation where the oldest ice preserves ‘long’ snapshots of mean atmospheric CO_2 , CH_4 , and $\delta^{18}\text{O}_{\text{atm}}$ over one or more glacial cycles. Although this interpretation of the Allan Hills ice core records limits our ability to comment on the amplitude of glacial-interglacial variability in these properties, we believe they provide robust records of mean atmospheric composition for CO_2 and CH_4 over the Pleistocene. Climate dependent differences in accumulation rates may introduce a small interglacial bias, shifting the mean higher than the true glacial-cycle average, however our analysis shows this difference is small (within 20 ppm and 50 ppb, respectively). This bias does not appear to change with sample age and therefore likely does not impact the long-term trend observed in the ‘long’ snapshot*

record. If a climate-dependent preservation bias increased across the Pleistocene with the growth of larger ice sheets and more extreme glacial minima, the long-term trend exhibited in our record may be dampened by ~10 ppm (see Methods). Alternative interpretations of the long-term trends in our record can be found in the supplement.

Line 23. <300 ppm doesn't seem to come from the text, at line 205, it says <280 ppm, and the uncorrected samples are all around 250 or less.

Our reference to <300 ppm comes from all existing CO₂ data in the early Pleistocene, including the collection of d13C-corrected CO₂ from the deepest samples. While most of our samples indicate CO₂ was closer to 250 ppm, we write <300 ppm as a conservative statement that includes the one d13C-corrected sample at 2.6 Ma that suggests CO₂ was 308 +/- 16 ppm.

The reference to <280 ppm in line 205 is about the Plio-Pleistocene Transition, noting that none of our samples from before the PPT indicate CO₂ was higher than 280 ppm. This is noteworthy considering it has been suggested that CO₂ falling below 280 ppm was a necessary ingredient for the buildup of land ice in the Northern Hemisphere at the start of the Pleistocene. However, we agree that our samples are all around 250 ppm before the PPT, and have therefore changed this line to the following:

New Line 188: *Due to the associated age uncertainty, only one sample is very likely from the Pliocene (⁴⁰Ar_{atm} age of 3.1 ± 0.3 Ma), with an estimated CO₂ of 245 ± 40 ppm. All samples (pristine and respiration-corrected) with a mean ⁴⁰Ar_{atm} age older than 2.6 Ma (n = 6) suggest CO₂ was 256 ± 26 ppm before the Plio-Pleistocene Transition, lower than indicated by boron-based proxies.*

Line 138: "Good agreement between the sampled synthetic dataset and the distribution of Allan Hills samples for both 'short' (Fig. 2b) and 'long' (Fig. 2c)". The agreement for CO₂ in Fig 2b (short snapshots) is poor. Again I emphasise the key conundrum here is that the actual data (yellow histogram in CO₂ Figs 2b and 2c) are so different. No amount of thinning and averaging can turn the data in 2b to that in 2c for CO₂, so what has happened?

We have adjusted the wording in the main text to distinguish between the good agreement for all other parameters and the short snapshot CO₂ record:

NEW LINE 145: *Finally, the much larger discrepancy between 'short' and 'long' snapshot means in CO₂ compared to CH₄ or δ¹⁸O_{atm} could suggest a small amount of excess CO₂. A number of processes have been observed to alter the CO₂ concentration in ice cores, such as the presence of melt layers, carbonate or silicate weathering, oxidation of organic carbon, and post-coring effects^{23–27}. Apart from silicate weathering and carbonate precipitation, most known processes increase the CO₂ concentration in bubbles leading to an overestimate of atmospheric CO₂. While in some cases CO₂ can be highly enriched (for example, due to oxidation of organic matter near the bedrock interface), in other cases alteration is subtle, with differences between ice cores reported in the 5-20 ppm range^{28–}*

³⁰. *If present, the amount of enriched CO₂ is too small (5-20 ppm) to correct for considering existing uncertainties in the isotopic endmembers of the CO₂ mixing model (see Methods).*

Line 162-4. A reader would never work out from reading the main text how you got to these numbers, or even what assumptions you have made. It needs a better explanation.

We now have included a clearer section in the supplement titled **Alternative Interpretations** (pasted below). While we think it is important to include these possibilities in the supplement, we have removed mention of them from the main text because we agree that including them was confusing to the reader.

Main Text:

NEW LINE 157: *In summary, although aspects of the paleoclimate information contained in Allan Hills ice cores remain poorly understood, our results support an interpretation where the oldest ice preserves ‘long’ snapshots of mean atmospheric CO₂, CH₄, and $\delta^{18}\text{O}_{\text{atm}}$ over one or more glacial cycles. Although this interpretation of the Allan Hills ice core records limits our ability to comment on the amplitude of glacial-interglacial variability in these properties, we believe they provide robust records of mean atmospheric composition for CO₂ and CH₄ over the Pleistocene. Climate dependent differences in accumulation rates may introduce a small interglacial bias, shifting the mean higher than the true glacial-cycle average, however our analysis shows this difference is small (within 20 ppm and 50 ppb, respectively). This bias does not appear to change with sample age and therefore likely does not impact the long-term trend observed in the ‘long’ snapshot record. If a climate-dependent preservation bias increased across the Pleistocene with the growth of larger ice sheets and more extreme glacial minima, the long-term trend exhibited in our record may be dampened by ~10 ppm (see Methods). Alternative interpretations of the long-term trends in our record can be found in the supplement.*

Supplementary Text:

NEW LINE S150:

Alternative interpretations

In the main text, we report the most probable interpretation of our ALHIC1901 record, assuming measurements from below 140 m represent mean values of one or more glacial cycles (‘long’ snapshots). Included below is a brief discussion of alternative approaches to interpreting our record.

Considering the sample number gets smaller with increasing age in the current dataset (late Pliocene samples: $n = 2$), we analyzed the ‘long’ snapshot record by binning samples in time and conducting t-tests assuming unequal variance. The mean CO₂ and CH₄ values were significantly different ($p\text{-value} < 0.05$) between the pre-MPT (>1.5 Ma, $n=10$) and post-MPT (<0.8 Ma, $n = 16$) samples (Figure S12), suggesting a drop in CO₂ and CH₄ of 10 ppm and 10 ppb between these two binned time periods, respectively.

We also must consider the possibility that the time resolution of each sample is the same, regardless of depth. This would be the case if the drastic increase in age density that we observe at 140 m is an artifact of large sections of the record being missing rather than a change in time resolution in a continuous but compressed record. While this seems an

unlikely scenario, to examine the implications we conducted a t-test between early Pleistocene (older than 1.5 Ma) and all late Pleistocene (400-800 ka) samples ('short' and 'long' snapshots). Results show lower mean CO₂ in the older samples and an insignificant difference in mean CH₄ (Figure S12). Under this endmember interpretation, mean CH₄ shows no change from 2.9-0.8 Ma while CO₂ exhibits a small, ~7 ppm increase.

Line 212-3. This is probably for the editors but it is not OK to use this unpublished data here, so unless the MOT paper is also accepted this sentence should be removed and just the marine core evidence used at this point.

We will address the issue of referencing the MOT paper upon request from the editors.

Line 239. I think you need to add a further caveat: the temperature could explain the CO₂ reduction but this is not only assuming sources and sinks in balance but also that ocean temperature change didn't affect other processes such as biological productivity, ocean circulation, etc.

Thank you for this feedback, we have edited our text to include this caveat

NEW LINE 220: On timescales of hundreds of thousands to millions of years, the global carbon cycle strikes a balance between the CO₂ emitted from Earth's interior and long-term carbon burial as organic matter and carbonate minerals⁴⁰. Earth's surface temperature is believed to play a critical role in this balance as a controlling factor on silicate weathering rates⁴¹. However, there is considerable debate as to the strength of the CO₂ thermostat as well as whether the controls to the thermostat reside on the continents or in the seafloor⁴² and whether or not changes in the global carbon cycle are driven primarily by changes in CO₂ fluxes or continental weatherability⁴³. Assuming long-term global natural CO₂ sources of ~0.1 Gt C/yr⁴⁴, our record of a ~20 ppm decline in atmospheric CO₂ over the Pleistocene indicates that sources and sinks of CO₂ have been in balance to within <1%⁴⁴.

There are three possible explanations for this tight balance; 1) it reflects the CO₂-thermostat responding to a decline in CO₂ sources; 2) it reflects the CO₂-thermostat responding to an increase in continental weatherability; 3) the CO₂-thermostat is weak and CO₂ sources and sinks are invariant and constant over the last 3 million years. Although we cannot distinguish between these hypotheses with existing data, estimates of the temperature-dependence of the silicate weathering feedback⁴² and our record of global cooling over the Pleistocene suggest that a ~5% decline in CO₂ sources (hypothesis #1) or similar increases in continental weatherability (hypothesis #2) are sufficient to explain the observed CO₂ decline. Hypothesis #3 cannot be excluded as the increase in CO₂ solubility associated with 1.5-2°C of global cooling would, all else remaining constant, lower atmospheric CO₂ by 12 ppm⁴⁵.

S124-5. This seems circular to me: you are saying there is not a change in the bias because the temperature does what you are expecting, but this is also a result of the paper.

Former Line 124-5: *Lastly, though a preservation bias toward glacial periods is not observed in the younger part of our record, we acknowledge that it is possible due to differing ice rheology leading to a change in ice flow between interglacial and glacial layers²¹. Evidence against this hypothesis comes from the expected cooling trend across the Pleistocene observed in both the water isotope record and mean ocean temperature record²² from the same ice core. In addition, we would expect to see a difference in $\delta^{38}/^{36}\text{Ar}$ if there was a switch in preservation bias from glacial layers in the early Pleistocene toward interglacial layers in the late Pleistocene.*

It is a fair point to say that we cannot rule out a bias toward glacial periods, however whatever bias may exist would impact all ice and gas records constructed from the core. The cooling trend observed in water isotopes and mean ocean temperature from the same samples where we observe stable greenhouse gas concentrations at least shows us that we are not missing the climatic trend expected for this time period. If the correlation between CO₂ and temperature that we observe in the late Pleistocene existed, we would expect a larger decline in CO₂ than what we have measured considering the water isotope and MOT record.

We have edited this sentence to hopefully make this point more clear:

NEW LINE S144: *Furthermore, whatever preservation bias may exist would impact all ice and gas records constructed from the ALHIC1901 core. The cooling trend across the Pleistocene observed in both the water isotope record¹⁶ and mean ocean temperature record²⁵ indicates the relatively stable concentrations of CO₂ and CH₄ from co-registered samples are not a result of the record missing a climatic trend that we would expect to see for this time period.*

Line S174 and following better explains why you have different scenarios, but again if it is your preferred scenarios you need to explain why “long-exposure snapshots in near-basal ice that preserve climate signals that more closely represent the mean over a glacial cycle” when you have said that there is an interglacial bias throughout.

The critical point is that changing preservation bias does not change the mean value of any of the three gas parameters by much (see above comments for more detail) and thus it makes sense that the long snapshots closely resemble EDC averages. This point is more clearly stated in our new version of the manuscript (and discussed in response to above comments).

Fig S4: y-axis labels are wrong (0.1, 0.2, 0.3 each appear twice).

The y-axis label has been fixed.

Remaining issues from referee 2

This third version of the paper is substantially revised from the last one, and does attempt to respond to all the points I (and the other reviewer) made on version 2. In doing so it takes a new tack: previously it noted the interglacial bias apparent in the short snapshots, but asserted that the long snapshots were some kind of average value of GHG concentrations. In the new version a substantial new section attempts to dismiss the short snapshots as somehow biased (either by being unrepresentative or through an addition of CO₂). This gives the authors a clear run to suggest that the long snapshots really are some kind of time average, and that this is likely to be unbiased through time. My feeling is that, had they decided on this route in version 1, much of the discussion of short snapshots might have gone into the supplement and we might have had a simpler and clearer paper, but I can see how we ended up where we are.

I think at this stage we have gone as far as we can in trying to unravel what the long snapshots really mean. I feel sure that everything I suggest the authors will already have thought of and discussed among themselves. The truth is that there are some mysteries in the data; we don't really understand the reason the short snapshots fall away from the long ones, and we don't really know what kind of average the long snapshots represent. However, the data deserve to be treated alongside the other major CO₂ datasets that are out there, and it's at least the simplest working hypothesis that whatever bias is present stays constant with time. Once we accept that this is the best that can be done at this stage the conclusions of the paper follow. I am therefore now recommending publication, although I do have a few suggestions about how the work is presented.

Thank you for these comments, we have really appreciated the feedback.

For me the critical point remains that the long snapshots are undoubtedly some kind of weighted average across the time they represent. The physics of higher snow accumulation in interglacials makes a weighting towards interglacials a starting point, but ice dynamical differential thinning might lead to other weightings. I would therefore recommend that this is explicitly stated at key points, ie that the values are referred to as weighted averages that may not be simple time averages. This then also allows the authors to explicitly state that they have made an assumption that the weighting across climate stages is invariant. I will suggest in my detailed comments where this might be done.

This is a good recommendation and one that we have implemented throughout the text. Wherever we discuss the long snapshot record as “mean” values, we have adjusted the wording to make it clear these are likely weighted averages. Related to this, when we compare our record to the overlapping EDC values from 500-800 ka, we state that we are comparing the time-averaged EDC value.

The key conclusion of the paper is that there is only a small CO₂ trend across the Pleistocene (250 to 230 ppm). The authors acknowledge that preservation bias could account for up to 20 ppm. They don't really consider the fact that a trend of this order could also be induced by changes in cycle length and that shorter cycles before the MPT might also account for a part of the trend. Take a hypothetical situation where glacials were 200 ppm and interglacials 250 ppm

throughout the Pleistocene. With 40 kyr cycles split 50:50 between glacial and interglacial and no preservation bias, the “average” would be 225 ppm, while with 100 kyr cycles split 80:20, the average would be 210 ppm, giving a 15 ppm trend that would be somewhat spurious. It would be useful to discuss this somewhere, even if using less simplistic calculations than I did.

I believe that the hypothetical situation you are suggesting is similar to the one we discuss in the supplement under “Climate-dependent preservation bias” and now have moved to the main text. Along the same lines as your calculations, we calculate an endmember scenario under which there was virtually no preservation bias in the late Pliocene, causing the oldest samples to be unweighted and the youngest samples to be the most weighted. In other words, the oldest samples would represent a 50:50 split between glacial and interglacial periods as you suggest. To calculate the CO₂ trend across the Pleistocene with this change in bias, we treated the late Pliocene samples as unweighted averages and compare this value to the time-averaged value of EDC from 500-800 ka. This gives us a 30 ppm decline, ~10 ppm larger than our observed trend.

Revised version of the Main text: *If, however, this bias did change, we would expect the difference in accumulation rates between glacial and interglacial layers to become more extreme with the growth of larger ice sheets and more extreme glacial minima. Thus, we may expect to see an increase in the weighting toward interglacial periods across the Pleistocene. In this case, values measured at 2.9-2.7 Ma would represent less biased long-term atmospheric averages. The most extreme version of this scenario would be if accumulation rates were constant in time, and under this scenario the late Pliocene/early Pleistocene results (250 ppm and 507 ppb for CO₂ and CH₄ respectively) would be true time averages. Comparing these to the time averages of the late Pleistocene from EDC (220 ppm and 503 ppb, respectively) would imply a CO₂ decline of 30 ppm, ~10 ppm larger than the observed trend in our ‘long’ snapshot record, and no significant change in CH₄.*

Detailed comments:

Line 17: I suggest “likely represent some kind of weighted averages over glacial cycles. Assuming the weighting stays constant, our values permit direct reconstructions of atmospheric greenhouse gases...”.

Completed

Line 129 “aye{Citation}rs” – check.

Completed

Line 119-155. While I understand this discussion it is a lot of text in the main paper, and quite peripheral to the main discussion as your conclusion from it is to ignore whatever the short snapshots are telling you. I don’t object to it being here but wonder whether there is material that would be more important in the main text than this?

We have moved some of this section to the supplement per this recommendation. The section in the supplement is titled “Short vs long snapshots”.

Line 131-132: Your suggestion about the use of ^{38}Ar doesn't really work, because you explicitly believe the short snapshots are mainly interglacial while the long snapshots are means, and yet Fig S9 shows an unchanged $^{38}/^{36}\text{Ar}$ value between short and long snapshots of the same age. This implies that this measure is not diagnostic of climate state. Unless you can explain this observation you need to downweight this evidence.

We are using $^{38}/^{36}\text{Ar}$ to say that whatever bias exists is constant through time. Our final conclusion is that all samples likely contain a weighting toward interglacial values, but that this is not the explanation for why short and long snapshots differ from one another. We explain this with Figure S5 that shows a large change in preservation bias does not shift the mean value of the gas distribution by much. Our final thoughts for the difference between short and long snapshots are either this is due to a sampling bias (and if we could sample more of the shallower section of the core we may find a larger range of values) or slight deviations in atmospheric CO_2 due to oxidation or organic carbon or some other process that is increasing the CO_2 by 5-20 ppm.

Line 132: you mean Fig S9 not Fig S8.

Completed

Line 158: I suggest a couple of additions (which I have placed in brackets): “where the oldest ice preserves CO_2 , CH_4 , and $\delta^{18}\text{O}_{\text{atm}}$ averaged (with an undetermined weighting between glacial and interglacial) over one or more glacial cycles. Although this interpretation of the Allan Hills ice core records limits our ability to comment on the amplitude of glacial-interglacial variability in these properties, we believe they provide robust records of (weighted-)mean atmospheric composition for CO_2 and CH_4 over long intervals...”

Completed, though with slightly different wording:

“In summary, although aspects of the paleoclimate information contained in Allan Hills ice cores remain poorly understood, our results support an interpretation where the oldest ice preserves CO_2 , CH_4 , and $\delta^{18}\text{O}_{\text{atm}}$ averaged over one or more glacial cycles with a weighting between glacial and interglacial periods that is relatively constant through time though not completely understood. Although this interpretation of the Allan Hills ice core records limits our ability to comment on the amplitude of glacial-interglacial variability in these properties, we believe they provide robust records of mean atmospheric composition for CO_2 and CH_4 over long intervals of the Pleistocene. This mean value is likely weighted due to climate dependent differences in accumulation rates that may introduce a small interglacial bias, shifting the mean CO_2 and CH_4 concentrations higher than the true glacial-cycle average.”

Line 166: “This bias does not appear to change with sample age”. While I agree this is your assumption I don't see on what basis you are asserting it to be a fact. Clearly we have no way of knowing whether interglacials are overrepresented in some samples and not others, and at some ages and not others. What your calculations show is that there can be a bias of up to 20 ppm, and we have no way of knowing which samples this applies to. Your reasonable assumption is that it stays the same at all ages, but have you really done anything that allows you to assert that it appears to not change with time?

This is a fair comment. We have adjusted the wording to emphasize that we can only speculate about preservation bias. The new section is pasted below:

Stable values of $\delta^{38/36}\text{Ar}$ (Extended Data Fig. 7b), an indicator of changes in firn column height, suggest that a preservation bias driven by accumulation rate is constant with sample age. If this is the case, the long-term trend observed in the ‘long’ snapshot record remains (~20 ppm decline from 2.8 – 1.2 Ma). If, however, this bias did change, we would expect the difference in accumulation rates between glacial and interglacial layers to become more extreme with the growth of larger ice sheets and more extreme glacial minima. Thus, we may expect to see an increase in the weighting toward interglacial periods across the Pleistocene. In this case, values measured at 2.9–2.7 Ma would represent less biased long-term atmospheric averages. The most extreme version of this scenario would be if accumulation rates were constant in time, and under this scenario the late Pliocene/early Pleistocene results (250 ppm and 507 ppb for CO_2 and CH_4 respectively) would be true time averages. Comparing these to the time averages of the late Pleistocene from EDC (220 ppm and 503 ppb, respectively) would imply a CO_2 decline of 30 ppm, ~10 ppm larger than the observed trend in our ‘long’ snapshot record, and no significant change in CH_4 . We describe alternative interpretations of the long-term trends in our record in the supplement.

~line 185. Would you like to add a sentence reflecting my comment about the proportion of time spent in glacial and interglacial states.? I think I can accept if you prefer not to add this complication.

See pasted section above – we have moved the wording about this hypothetical change in weighting that was in the supplement into this part of the main text and hope it satisfies your concern.

Line 188. I wonder if it would be worth adding that your values fall within the range of 11B data, but for them to be compatible with 11B-based reconstructions, your samples would have to have a massive bias towards glacial ice, which seems improbable.

Completed

Line 247: I suggest “likely reflecting atmospheric compositions averaged with some kind of climate weighting over glacial cycles”. (Since in the next sentence you say they have a bias they are clearly not a mean in the sense most readers would envisage it). And then “a reasonable assumption is that this bias does not change with sample age”.

Completed

Ref 13 is incomplete but i believe is now in press.

Completed

Supplementary. The discussion of preservation bias that leads to Figs S7, S8 and S5 is very hard

to follow. I just about grasped it but it is not straightforward. Please take one last look to see if you can explain it more simply: a start might be to reorder the figures so that we can see the scenarios you are calculating (S7, S8) from before we see the result in S5!

We changed the figure numbers as suggested, and have added some more detail to this section in the supplement. Hopefully it is clearer! Parts that have been edited are pasted below:

To explore the effects of a climate-dependent preservation bias, we created a synthetic ice core record that has glacial-to-interglacial annual layer thickness variations we would expect at Allan Hills. We limited the synthetic record to 400-800 ka to overlap with ALHIC1901 measurements within $^{40}\text{Ar}_{\text{atm}}$ age uncertainty. We first calculated the annual layer thicknesses of EDC and the S27 ice core from Allan Hills using their respective chronologies^{16,21} (after evenly spacing the records with respect to time). Comparing the two records of annual layer thickness revealed an order-of-magnitude difference (Extended Data Fig. 5b). We thus exaggerated the EDC layer thickness pattern to produce a synthetic depth record for EDC with a more extreme difference between glacial and interglacial layers. This was done by first identifying the peaks and troughs of the original EDC layer thickness record and interpolating between them using the nearest neighbor method to produce a simplified layer thickness record with a square wave pattern (Extended Data Fig. 5a). We varied the amplitude of the simplified square-wave record, increasing the relative interglacial-to-glacial layer thickness contrast by a factor of 10 to better match the layer thickness record of the S27 ice core. This was done by multiplying the peak and trough values of the simplified record by 5 and 0.5 respectively (Extended Data Fig. 5b). We then calculated synthetic depths using the new layer thickness record and subsequently interpolated the EDC gas records^{20,21,29,46,47}, producing hypothetical continuous records with more extreme weighting toward interglacial layers. We also include the results of a layer thickness contrast 3 orders of magnitude larger than the simplified square-wave record (multiplying the peak and trough values by 50 and 0.05, respectively), providing an extreme preservation bias scenario (Extended Data Fig. 5b). The difference in distributions for each gas parameter between the simplified EDC layer thickness record and the two exaggerated records is shown in Extended Data Fig. 6. An interglacial bias is observed in all three distributions relative to the unweighted average. Considering our 'long' snapshots are likely average values of glacial cycles, the difference in the means of each distribution is most informative. We find that the mean value for each gas property does not change very much with increased weighting. For the single order of magnitude scenario, mean CO_2 was 11 ppm higher than the unweighted average and CH_4 was 34 ppb higher. Even under the extreme (and unreasonable) preservation bias scenario that is 3 orders of magnitude more exaggerated than EDC, mean CO_2 and CH_4 are only 20 ppm and 50 ppb higher than the unweighted average, respectively.

Thus, our analysis suggests that even with an order of magnitude more extreme change in layer thickness between glacials and interglacials (as we might expect in a place like the Allan Hills BIA), the absolute concentrations are within ~10 ppm and ~30 ppb of the true average. Furthermore, if the preservation bias is constant through our record, such weighting does not impact the long-term trend. This leads to the question of whether the preservation bias at the Allan Hills BIA has changed across the Pleistocene.

Evidence for a change in preservation bias might be observed in the $\delta^{38/36}\text{Ar}$ ratio, which records changes in the height of the diffusive zone within the firn column via gravitational enrichment. Most if not all polar ice cores show a glacial-interglacial difference in firn thickness. The value of $\delta^{38/36}\text{Ar}$ is relatively constant throughout our whole record, indicating little change in the gradient between glacial and interglacial layer thickness (Extended Data Fig. 7b). The stability in $\delta^{38/36}\text{Ar}$ would be consistent with, but does not prove, a preservation bias that is constant throughout the record...